

# A review of the Statistical Mechanics approach to Random Optimization Problems

Fabrizio Altarelli<sup>1,2</sup>, Rémi Monasson<sup>2</sup>, Guilhem Semerjian<sup>2</sup> and Francesco Zamponi<sup>2</sup>

<sup>1</sup> *Dipartimento di Fisica and CNR, Università di Roma La Sapienza, P. A. Moro 2, 00185 Roma, Italy,*

<sup>2</sup> *LPTENS, Unité Mixte de Recherche (UMR 8549) du CNRS et de l'ENS, associée à l'UPMC Univ Paris 06, 24 Rue Lhomond, 75231 Paris Cedex 05, France.*

We review the connection between statistical mechanics and the analysis of random optimization problems, with particular emphasis on the random  $k$ -SAT problem. We discuss and characterize the different phase transitions that are met in these problems, starting from basic concepts. We also discuss how statistical mechanics methods can be used to investigate the behavior of local search and decimation based algorithms.

*This paper has been written as a contribution to the “Handbook of Satisfiability” to be published in 2008 by IOS press.*

## I. INTRODUCTION

The connection between the statistical physics of disordered systems and optimization problems in computer science dates back from twenty years at least [1]. In combinatorial optimization one is given a cost function (the length of a tour in the traveling salesman problem (TSP), the number of violated constraints in constraint satisfaction problems, ...) over a set of variables and looks for the minimal cost over an allowed range for those variables. Finding the true minimum may be complicated, and requires bigger and bigger computational efforts as the number of variables to be minimized over increases [2]. Statistical physics is at first sight very different. The scope is to deduce the macroscopic, that is, global properties of a physical system, for instance a gas, a liquid or a solid, from the knowledge of the energetic interactions of its elementary components (molecules, atoms or ions). However, at very low temperature, these elementary components are essentially forced to occupy the spatial conformation minimizing the global energy of the system. Hence low temperature statistical physics can be seen as the search for minimizing a cost function whose expression reflects the laws of Nature or, more humbly, the degree of accuracy retained in its description. This problem is generally not difficult to solve for non disordered systems where the lowest energy conformation are crystals in which components are regularly spaced from each other. Yet the presence of disorder, e.g. impurities, makes the problem very difficult and finding the conformation with minimal energy is a true optimization problem.

At the beginning of the eighties, following the works of G. Parisi and others on systems called spin glasses [1], important progresses were made in the statistical physics of disordered systems. Those progresses made possible the quantitative study of the properties of systems given some distribution of the disorder (for instance the location of impurities) such as the average minimal energy and its fluctuations. The application to optimization problems was natural and led to beautiful studies on (among others) the average properties of the minimal tour length in the TSP, the minimal cost in Bipartite Matching, for some specific instance distributions [1]. Unfortunately statistical physicists and computer scientists did not establish close ties on a large scale at that time. The reason could have been of methodological nature [3]. While physicists were making statistical statements, true for a given distribution of inputs, computer scientists were rather interested in solving one (or several) particular instances of a problem. The focus was thus on efficient ways to do so, that is, requiring a computational effort growing not too quickly with the number of data defining the instance. Knowing precisely the typical properties for a given, academic distribution of instances did not help much to solve practical cases.

At the beginning of the nineties practitioners in artificial intelligence realized that classes of random constraint satisfaction problems used as artificial benchmarks for search algorithms exhibited abrupt changes of behaviour when some control parameter were finely tuned [4]. The most celebrated example was random  $k$ -Satisfiability, where one looks for a solution to a set of random logical constraints over a set of Boolean variables. It appeared that, for large sets of variables, there was a critical value of the number of constraints per variable below which there almost surely existed solutions, and above which solutions were absent. An important feature was that the performances of known search algorithms drastically worsened in the vicinity of this critical ratio. In addition to its intrinsic mathematical interest the random  $k$ -SAT problem was therefore worth to be studied for ‘practical’ reasons.

This critical phenomenon, strongly reminiscent of phase transitions in condensed matter physics, led to a revival of the research at the interface between statistical physics and computer science, which is still very active. The purpose of the present review is to introduce the non physicist reader to some concepts required to understand the literature in the field and to present some major results. We shall in particular discuss the refined picture of the satisfiable phase put forward in statistical mechanics studies and the algorithmic approach (Survey Propagation, an extension of Belief Propagation used in communication theory and statistical inference) this picture suggested.

While the presentation will mostly focus on the  $k$ -Satisfiability problem (with random constraints) we will occasionally discuss another computational problem, namely, linear systems of Boolean equations. A good reason to do so is that this problem exhibits some essential features encountered in random  $k$ -Satisfiability, while being technically simpler to study. In addition it is closely related to error-correcting codes in communication theory.

The chapter is divided into four main parts. In Section II we present the basic statistical physics concepts necessary to understand the onset of phase transitions, and to characterize the nature of the phases. Those are illustrated on a simple example of decision problem, the so-called perceptron problem. In Section III we review the scenario of the various phase transitions taking place in random  $k$ -SAT. Section IV and V present the techniques used to study various type of algorithms in optimization (local search, backtracking procedures, message passing algorithms). We end up with some conclusive remarks in Sec. VI.

## II. PHASE TRANSITIONS: BASIC CONCEPTS AND ILLUSTRATION

### A. A simple decision problem with a phase transition: the continuous perceptron

For pedagogical reasons we first discuss a simple example exhibiting several important features we shall define more formally in the next subsection. Consider  $M$  points  $\underline{T}^1, \dots, \underline{T}^M$  of the  $N$ -dimensional space  $\mathbb{R}^N$ , their coordinates being denoted  $\underline{T}^a = (T_1^a, \dots, T_N^a)$ . The continuous perceptron problem consists in deciding the existence of a vector  $\underline{\sigma} \in \mathbb{R}^N$  which has a positive scalar product with all vectors linking the origin of  $\mathbb{R}^N$  to the  $\underline{T}$ 's,

$$\underline{\sigma} \cdot \underline{T}^a \equiv \sum_{i=1}^N \sigma_i T_i^a > 0, \quad \forall a = 1, \dots, M, \quad (1)$$

or in other words determining whether the  $M$  points belong to the same half-space. The term continuous in the name of the problem emphasizes the domain  $\mathbb{R}^N$  of the variable  $\underline{\sigma}$ . This makes the problem polynomial from worst-case complexity point of view [5].

Suppose now that the points are chosen independently, identically, uniformly on the unit hypersphere, and call

$$P(N, M) = \text{Probability that a set of } M \text{ randomly chosen points} \\ \text{belong to the same half-space.}$$

This quantity can be computed exactly [6] (see also Chapter 5.7 of [5]) and is plotted in Fig. 1 as a function of the ratio  $\alpha = M/N$  for increasing sizes  $N = 5, 20, 100$ . Obviously  $P$  is a decreasing function of the number  $M$  of points for a given size  $N$ : increasing the number of constraints can only make more difficult the simultaneous satisfaction of all of them. More surprisingly, the figure suggests that, in the large size limit  $N \rightarrow \infty$ , the probability  $P$  reaches a limiting value 0 or 1 depending on whether the ratio  $\alpha$  lies, respectively, above or below some ‘critical’ value  $\alpha_s = 2$ . This is confirmed by the analytical expression of  $P$  obtained in [6],

$$P(N, M) = \frac{1}{2^{M-1}} \sum_{i=0}^{\min(N-1, M-1)} \binom{M-1}{i}, \quad (2)$$

from which one can easily show that, indeed,

$$\lim_{N \rightarrow \infty} P(N, M = N\alpha) = \begin{cases} 1 & \text{if } \alpha < \alpha_s \\ 0 & \text{if } \alpha > \alpha_s \end{cases}, \quad \text{with } \alpha_s = 2. \quad (3)$$

Actually the analytical expression of  $P$  allows to describe more accurately the drop in the probability as  $\alpha$  increases. To this aim we make a zoom on the transition region  $M \approx N\alpha_s$  and find from (2) that

$$\lim_{N \rightarrow \infty} P(N, M = N\alpha_s(1 + \lambda N^{-1/2})) = \int_{\lambda\sqrt{2}}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-x^2/2}. \quad (4)$$

As it should the limits  $\lambda \rightarrow \pm\infty$  gives back the coarse description of Eq. (3)

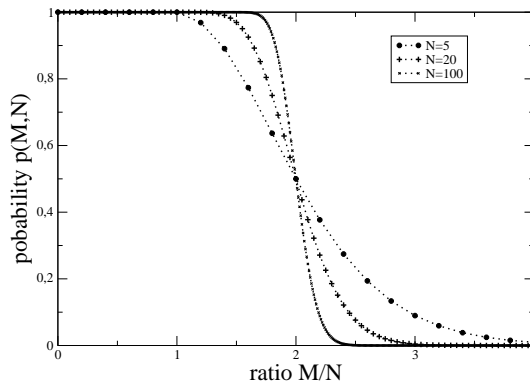


FIG. 1: Probability  $P(N, M)$  that  $M$  random points on the  $N$ -dimensional unit hypersphere are located in the same half-space. Symbols correspond to Cover's exact result [6], see Eq. (2), lines serve as guides to the eye.

## B. Generic definitions

We now put this simple example in a broader perspective and introduce some generic concepts that it illustrates, along with the definitions of the problems studied in the following.

- Constraint Satisfaction Problem (CSP)

A CSP is a decision problem where an assignment (or configuration) of  $N$  variables  $\underline{\sigma} = (\sigma_1, \dots, \sigma_N) \in \mathcal{X}^N$  is required to simultaneously satisfy  $M$  constraints. In the continuous perceptron the domain of  $\underline{\sigma}$  is  $\mathbb{R}^N$  and the constraints impose the positivity of the scalar products (1). The instance of the CSP, also called formula in the following, is said satisfiable if there exists a solution (an assignment of  $\underline{\sigma}$  fulfilling all the constraints). The  $k$ -SAT problem is a boolean CSP ( $\mathcal{X} = \{\text{True}, \text{False}\}$ ) where each constraint (clause) is the disjunction (logical OR) of  $k$  literals (a variable or its negation). Similarly in  $k$ -XORSAT the literals are combined by an eXclusive OR operation, or equivalently an addition modulo 2 of 0/1 boolean variables is required to take a given value. The worst-case complexities of these two problems are very different ( $k$ -XORSAT is in the P complexity class for any  $k$  while  $k$ -SAT is NP-complete for any  $k \geq 3$ ), yet for the issues of this review we shall see that they present a lot of similarities. In the following we use the statistical mechanics convention and represent boolean variables by Ising spins,  $\mathcal{X} = \{-1, +1\}$ . A  $k$ -SAT clause will be defined by  $k$  indices  $i_1, \dots, i_k \in [1, N]$  and  $k$  values  $J_{i_1}, \dots, J_{i_k} = \pm 1$ , such that the clause is unsatisfied by the assignment  $\underline{\sigma}$  if and only if  $\sigma_{i_j} = J_{i_j} \quad \forall j \in [1, k]$ . A  $k$ -XORSAT clause is satisfied if the product of the spins is equal to a fixed value,  $\sigma_{i_1} \dots \sigma_{i_k} = J$ .

- random Constraint Satisfaction Problem (rCSP)

The set of instances of most CSP can be turned in a probabilistic space by defining a distribution over its constraints, as was done in the perceptron case by drawing the vertices  $\underline{T}^a$  uniformly on the hypersphere. The random  $k$ -SAT formulas considered in the following are obtained by choosing for each clause  $a$  independently a  $k$ -uplet of distinct indices  $i_1^a, \dots, i_k^a$  uniformly over the  $\binom{N}{k}$  possible ones, and negating or not the corresponding literals ( $J_i^a = \pm 1$ ) with equal probability one-half. The indices of random XORSAT formulas are chosen similarly, with the constant  $J^a = \pm 1$  uniformly.

- thermodynamic limit and phase transitions

These two terms are the physics jargon for, respectively, the large size limit ( $N \rightarrow \infty$ ) and for threshold phenomena as stated for instance in (3). In the thermodynamic limit the typical behavior of physical systems is controlled by a small number of parameters, for instance the temperature and pressure of a gas. At a phase transition these systems are drastically altered by a tiny change of a control parameter, think for instance at what happens to water when its temperature crosses 100 °C. This critical value of the temperature separates two qualitatively distinct phases, liquid and gaseous. For random CSPs the role of control parameter is usually played by the ratio of constraints per variable,  $\alpha = M/N$ , kept constant in the thermodynamic limit. Eq. (3) describes a satisfiability transition for the continuous perceptron, the critical value  $\alpha_s = 2$  separating a satisfiable phase at low  $\alpha$  where instances typically have solutions to a phase where they typically do not. Typically is used here as a synonym for with high probability, i.e. with a probability which goes to one in the thermodynamic limit.

- Finite Size Scaling (FSS)

The refined description of the neighborhood of the critical value of  $\alpha$  provided by (4) is known as a finite size scaling relation. More generally the finite size scaling hypothesis for a threshold phenomenon takes the form

$$\lim_{N \rightarrow \infty} P(N, M = N\alpha_s(1 + \lambda N^{-1/\nu})) = \mathcal{F}(\lambda) , \quad (5)$$

where  $\nu$  is called the FSS exponent (2 for the continuous perceptron) and the scaling function  $\mathcal{F}(\lambda)$  has limits 1 and 0 at respectively  $-\infty$  and  $+\infty$ . This means that, for a large but finite size  $N$ , the transition window for the values of  $M/N$  where the probability drops from  $1 - \epsilon$  down to  $\epsilon$  is, for arbitrary small  $\epsilon$ , of width  $N^{-1/\nu}$ . Results of this flavour are familiar in the study of random graphs [7]; for instance the appearance of a giant component containing a finite fraction of the vertices of an Erdős-Rényi random graph happens on a window of width  $N^{-1/3}$  on the average connectivity. FSS relations are important, not only from the theoretical point of view, but also for practical applications. Indeed numerical experiments are always performed on finite-size instances while theoretical predictions on phase transitions are usually true in the  $N \rightarrow \infty$  limit. Finite-size scaling relations help to bridge the gap between the two. We shall review some FSS results in Sec. III E.

Let us emphasize that random  $k$ -SAT, and other random CSP, are expected to share some features of the continuous perceptron model, for instance the existence of a satisfiability threshold, but of course not its extreme analytical simplicity. In fact, despite an intensive research activity, the mere existence of a satisfiability threshold for random SAT formulas remains a (widely accepted) conjecture. A significant achievement towards the resolution of the conjecture was the proof by Friedgut of the existence of a non-uniform sharp threshold [8]. There exists also upper [9] and lower [10] bounds on the possible location of this putative threshold, which become almost tight for large values of  $k$  [11]. We refer the reader to the chapter [12] of this volume for more details on these issues. This difficulty to obtain tight results with the currently available rigorous techniques is a motivation for the use of heuristic statistical mechanics methods, that provide intuitions on why the standard mathematical ones run into trouble and how to amend them. In the recent years important results first conjectured by physicists were indeed rigorously proven. Before describing in some generality the statistical mechanics approach, it is instructive to study a simple variation of the perceptron model for which the basic probabilistic techniques become inefficient.

### C. The perceptron problem continued: binary variables

The binary perceptron problem consists in looking for solutions of (1) on the hypercube i.e. the domain of the variable  $\underline{\sigma}$  is  $\mathcal{X}^N = \{-1, +1\}^N$  instead of  $\mathbb{R}^N$ . This decision problem is NP-complete. Unfortunately Cover's calculation [6] cannot be extended to this case, though it is natural to expect a similar satisfiability threshold phenomenon at an a priori distinct value  $\alpha_s$ . Let us first try to study this point with basic probabilistic tools, namely the first and second moment method [13]. The former is an application of the Markov inequality,

$$\text{Prob}[Z > 0] \leq \mathbb{E}[Z] , \quad (6)$$

valid for positive integer valued random variables  $Z$ . We shall use it taking for  $Z$  the number of solutions of (1),

$$Z = \sum_{\underline{\sigma} \in \mathcal{X}^N} \prod_{a=1}^M \theta(\underline{\sigma} \cdot \underline{T}^a) , \quad (7)$$

where  $\theta(x) = 1$  if  $x > 0$ , 0 if  $x \leq 0$ . The expectation value of the number of solutions is easily computed,

$$\mathbb{E}[Z] = 2^N \times 2^{-M} = e^{N G_1} \quad \text{with} \quad G_1 = (1 - \alpha) \ln 2 , \quad (8)$$

and vanishes when  $N \rightarrow \infty$  if  $\alpha > 1$ . Hence, from Markov's inequality (6), with high probability constraints (1) have no solution on the hypercube when the ratio  $\alpha$  exceeds unity: if the threshold  $\alpha_s$  exists, it must satisfy the bound  $\alpha_s \leq 1$ . One can look for a lower bound to  $\alpha_s$  using the second moment method, relying on the inequality [13]

$$\frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]} \leq \text{Prob}[Z > 0] . \quad (9)$$

The expectation value of the squared number of solutions reads

$$\mathbb{E}[Z^2] = \sum_{\underline{\sigma}, \underline{\sigma}'} (\mathbb{E}[\theta(\underline{\sigma} \cdot \underline{T}) \theta(\underline{\sigma}' \cdot \underline{T})])^M \quad (10)$$

since the vertices  $\underline{T}^a$  are chosen independently of each other. The expectation value on the right hand side of the above expression is simply the probability that the vector pointing to a randomly chosen vertex,  $\underline{T}$ , has positive scalar product with both vectors  $\underline{\sigma}, \underline{\sigma}'$ . Elementary geometrical considerations reveal that

$$\mathbb{E}[\theta(\underline{\sigma} \cdot \underline{T}) \theta(\underline{\sigma}' \cdot \underline{T})] = \frac{1}{2\pi} (\pi - \varphi(\underline{\sigma}, \underline{\sigma}')) \quad (11)$$

where  $\varphi$  is the relative angle between the two vectors. This angle can be alternatively parametrized by the overlap between  $\underline{\sigma}$  and  $\underline{\sigma}'$ , i.e. the normalized scalar product,

$$q = \frac{1}{N} \sum_{i=1}^N \sigma_i \sigma'_i = 1 - 2 \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\sigma_i \neq \sigma'_i) . \quad (12)$$

The last expression, in which  $\mathbb{I}(E)$  denotes the indicator function of the event  $E$ , reveals the traduction between the concept of overlap and the more traditional Hamming distance. The sum over vectors in (10) can then be replaced by a sum over overlap values with appropriate combinatorial coefficients counting the number of pairs of vectors at a given overlap. The outcome is

$$\mathbb{E}[Z^2] = 2^N \sum_{q=-1, -1+\frac{2}{N}, -1+\frac{4}{N}, \dots, 1} \binom{N}{N(\frac{1+q}{2})} \left( \frac{1}{2} - \frac{1}{2\pi} \text{Arcos } q \right)^M . \quad (13)$$

In the large  $N$  limit we can estimate this sum with the Laplace method,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{E}[Z^2] = \max_{-1 < q < 1} G_2(q) , \quad (14)$$

where

$$\begin{aligned} G_2(q) = \ln 2 & - \left( \frac{1+q}{2} \right) \ln \left( \frac{1+q}{2} \right) - \left( \frac{1-q}{2} \right) \ln \left( \frac{1-q}{2} \right) \\ & + \alpha \ln \left( \frac{1}{2} - \frac{1}{2\pi} \text{Arcos } q \right) . \end{aligned} \quad (15)$$

Two conclusions can be drawn from the above calculation:

- no useful lower bound to  $\alpha_s$  can be obtained from such a direct application of the second moment method. Indeed, maximization of  $G_2$  (15) over  $q$  shows that  $\mathbb{E}[Z^2] \gg (\mathbb{E}[Z])^2$  when  $N$  diverges, whenever  $\alpha > 0$ , and in consequence the left hand side of (9) vanishes. A possible scenario which explains this absence of concentration of the number of solutions is the following. As shown by the moment calculation the natural scaling of  $Z$  is exponentially large in  $N$  (as is the total configuration space  $\mathcal{X}^N$ ). We shall thus denote  $s = (\ln Z)/N$  the random variable of order one counting the log degeneracy of the solutions. Suppose  $s$  follows a large deviation principle [14] that we state in a very rough way as  $\text{Prob}[s] \approx \exp[NL(s)]$ , with  $L(s)$  a negative rate function, assumed for simplicity to be concave. Then the moments of  $Z$  are given, at the leading exponential order, by

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{E}[Z^n] = \max_s [L(s) + ns] , \quad (16)$$

and are controlled by the values of  $s$  such that  $L'(s) = -n$ . The moments of larger and larger order  $n$  are thus dominated by the contribution of rarer and rarer instances with larger and larger numbers of solutions. On the contrary the typical value of the number of solutions is given by the maximum of  $L$ , reached in a value we denote  $s_g(\alpha)$ : with high probability when  $N \rightarrow \infty$ ,  $Z$  is comprised between  $e^{N(s_g(\alpha)-\epsilon)}$  and  $e^{N(s_g(\alpha)+\epsilon)}$ , for any  $\epsilon > 0$ . From this reasoning it appears that the relevant quantity to be computed is

$$s_g(\alpha) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[\ln Z] = \lim_{N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{1}{n} \ln \mathbb{E}[Z^n] . \quad (17)$$

This idea of computing moments of vanishing order is known in statistical mechanics as the replica<sup>1</sup> method [1]. Its non-rigorous implementation consists in determining the moments of integer order  $n$ , which are then continued

---

<sup>1</sup> The vocable replicas comes from the presence of  $n$  copies of the vector  $\underline{\sigma}$  in the calculation of  $Z^n$  (see the  $n = 2$  case in formula (10)).

towards  $n = 0$ . The outcome of such a computation for the binary perceptron problem reads [15]

$$s_g(\alpha) = \max_{q, \hat{q}} \left\{ -\frac{1}{2}q(1-\hat{q}) + \int_{-\infty}^{\infty} Dz \ln(2 \cosh(z\sqrt{\hat{q}})) + \alpha \int_{-\infty}^{\infty} Dz \ln \left[ \int_{z\sqrt{q/(1-q)}}^{\infty} Dy \right] \right\}, \quad (18)$$

where  $Dz \equiv dz e^{-z^2/2}/\sqrt{2\pi}$ . The entropy  $s_g(\alpha)$  is a decreasing function of  $\alpha$ , which vanishes in  $\alpha_s \simeq 0.833$ . Numerical experiments support this value for the critical ratio of the satisfiable/unsatisfiable phase transition.

- the calculation of the second moment is naturally related to the determination of the value of the overlap  $q$  between pairs of solutions (or equivalently their Hamming distance, recall Eq. (12)). This conclusion extends to the calculation of the  $n^{th}$  moment for any integer  $n$ , and to the  $n \rightarrow 0$  limit. The value of  $q$  maximizing the r.h.s. of (18),  $q^*(\alpha)$ , represents the average overlap between two solutions of the same set of constraints (1). Actually the distribution of overlaps is highly concentrated in the large  $N$  limit around  $q^*(\alpha)$ , in other words the (reduced) Hamming distance between two solutions is, with high probability, equal to  $d^*(\alpha) = (1 - q^*(\alpha))/2$ . This distance  $d^*(\alpha)$  ranges from  $\frac{1}{2}$  for  $\alpha = 0$  to  $\simeq \frac{1}{4}$  at  $\alpha = \alpha_s$ . Slightly below the critical ratio solutions are still far away from each other on the hypercube<sup>2</sup>.

Note that the perceptron problem is not as far as it could seem from the main subject of this review. There exists indeed a natural mapping between the binary perceptron problem and  $k$ -SAT. Assume the vertices  $\underline{T}$  of the perceptron problem, instead of being drawn on the hypersphere, have coordinates that can take three values:  $T_i = -1, 0, 1$ . Consider now a  $k$ -SAT formula  $F$ . To each clause  $a$  of  $F$  we associate the vertex  $\underline{T}^a$  with coordinates  $T_i^a = -J_i^a$  if variable  $i$  appears in clause  $a$ , 0 otherwise. Of course  $\sum_i |T_i^a| = k$ : exactly  $k$  coordinates have non zero values for each vertex. Then replace condition (1) with

$$\sum_{i=1}^N \sigma_i T_i^a > -(k-1), \quad \forall a = 1, \dots, M. \quad (19)$$

The scalar product is not required to be positive any longer, but to be larger than  $-(k-1)$ . It is an easy check that the perceptron problem admits a solution on the hypercube ( $\sigma_i = \pm 1$ ) if and only if  $F$  is satisfiable. While in the binary perceptron model all coordinates are non-vanishing, only a finite number of them take non zero values in  $k$ -SAT. For this reason  $k$ -SAT is called a diluted model in statistical physics.

Also the direct application of the second moment method fails for the random  $k$ -SAT problem; yet a refined version of it was used in [11], which leads to asymptotically (at large  $k$ ) tight bounds on the location of the satisfiability threshold.

#### D. From random CSP to statistical mechanics of disordered systems

The binary perceptron example taught us that the number of solutions  $Z$  of a satisfiable random CSP usually scales exponentially with the size of the problem, with large fluctuations that prevent the direct use of standard moment methods. This led us to the introduction of the quenched entropy, as defined in (17). The computation techniques used to obtain (18) were in fact developed in an apparently different field, the statistical mechanics of disordered systems [1].

Let us review some basic concepts of statistical mechanics (for introductory books see for example [16, 17]). A physical system can be modeled by a space of configuration  $\underline{\sigma} \in \mathcal{X}^N$ , on which is defined an energy function  $E(\underline{\sigma})$ . For instance usual magnets are described by Ising spins  $\sigma_i = \pm 1$ , the energy being minimized when adjacent spins take the same value. The equilibrium properties of a physical system at temperature  $T$  are given by the Gibbs-Boltzmann probability measure on  $\mathcal{X}^N$ ,

$$\mu(\underline{\sigma}) = \frac{1}{Z} \exp[-\beta E(\underline{\sigma})], \quad (20)$$

---

<sup>2</sup> This situation is very different from the continuous perceptron case, where the typical overlap  $q^*(\alpha)$  reaches one when  $\alpha$  tends to 2: a single solution is left right at the critical ratio.

where the inverse temperature  $\beta$  equals  $1/T$  and  $Z$  is a normalization called partition function. The energy function  $E$  has a natural scaling, linear in the number  $N$  of variables (such a quantity is said to be extensive). In consequence in the thermodynamic limit the Gibbs-Boltzmann measure concentrates on configurations with a given energy density ( $e = E/N$ ), which depends on the conjugated parameter  $\beta$ . The number of such configurations is usually exponentially large,  $\approx \exp[Ns]$ , with  $s$  called the entropy density. The partition function is thus dominated by the contribution of these configurations, hence  $\lim(\ln Z/N) = s - \beta e$ .

In the above presentation we supposed the energy to be a simple, known function of the configurations. In fact some magnetic compounds, called spin-glasses, are intrinsically disordered on a microscopic scale. This means that there is no hope in describing exactly their microscopic details, but that one should rather assume their energy to be itself a random function with a known distribution. Hopefully in the thermodynamic limit the fluctuations of the thermodynamic observables as the energy and entropy density vanish, hence the properties of a typical sample will be closely described by the average (over the distribution of the energy function) of the entropy and energy density.

The random CSPs fit naturally in this line of research. The energy function  $E(\underline{\sigma})$  of a CSP is defined as the number of constraints violated by the assignment  $\underline{\sigma}$ , in other words this is the cost function to be minimized in the associated optimization problem (MAXSAT for instance). Moreover the distribution of random instances of CSP is the counterpart of the distribution over the microscopic description of a disordered solid. The study of the optimal configurations of a CSP, and in particular the characterization of a satisfiability phase transition, is achieved by taking the  $\beta \rightarrow \infty$  limit. Indeed, when this parameter increases (or equivalently the temperature goes to 0), the law (20) favors the lowest energy configurations. In particular if the formula is satisfiable  $\mu$  becomes the uniform measures over the solutions. Two important features of the formula can be deduced from the behavior of  $Z$  at large  $\beta$ : the ground-state energy  $E_g = \min_{\underline{\sigma}} E(\underline{\sigma})$ , which indicates how good are the optimal configurations, and the ground state entropy  $S_g = \ln(|\{\underline{\sigma} : E(\underline{\sigma}) = E_g\}|)$ , which counts the degeneracy of these optimal configurations. The satisfiability of a formula is equivalent to its ground-state energy being equal to 0. In the large  $N$  limit these two thermodynamic quantities are supposed to concentrate around their mean values (this is proven for  $E$  in [18]), we thus introduce the associated typical densities,

$$e_g(\alpha) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[E_g], \quad s_g(\alpha) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[S_g]. \quad (21)$$

Notice that formula (21) coincides with (17) in the satisfiable phase (where the ground state energy vanishes).

Some criteria are needed to relate these thermodynamic quantities to the (presumed to exist) satisfiability threshold  $\alpha_s$ . A first approach, used for instance in [19], consists in locating it as the point where the ground-state energy density  $e_g$  becomes positive. The assumption underlying this reasoning is the absence of an intermediate, typically UNSAT regime, with a sub-extensive positive  $E_g$ . In the discussion of the binary perceptron we used another criterion, namely we recognized  $\alpha_s$  by the cancellation of the ground-state entropy density. This argument will be true if the typical number of solutions vanishes continuously at  $\alpha_s$ . It is easy to realize that this is not the case for random  $k$ -SAT: at any finite value of  $\alpha$  a finite fraction  $\exp[-\alpha k]$  of the variables do not appear in any clause, which leads to a trivial lower bound  $(\ln 2) \exp[-\alpha k]$  on  $s_g$ . This quantity is thus finite at the transition, a large number of solutions disappear suddenly at  $\alpha_s$ . Even if it is wrong, the criterion  $s_g(\alpha) = 0$  for the determination of the satisfiability transition is instructive for two reasons. First, it becomes asymptotically correct at large  $k$  (free variables are very rare in this limit), this is why it works for the binary perceptron of Section II C (which is, as we have seen, close to  $k$ -SAT with  $k$  of order  $N$ ). Second, it will reappear below in a refined version: we shall indeed decompose the entropy in two qualitatively distinct contributions, one of the two being indeed vanishing at the satisfiability transition.

### III. PHASE TRANSITIONS IN RANDOM CSPS

#### A. The clustering phenomenon

We have seen that the statistical physics approach to the perceptron problem naturally provided us with information about the geometry of the space of its solutions. Maybe one of the most important contribution of physicists to the field of random CSP was to suggest the presence of further phase transitions in the satisfiable regime  $\alpha < \alpha_s$ , affecting qualitatively the geometry (structure) of the set of solutions [20–22].

This subset of the configuration space is indeed thought to break down into “clusters” in a part of the satisfiable phase,  $\alpha \in [\alpha_d, \alpha_s]$ ,  $\alpha_d$  being the threshold value for the clustering transition. Clusters are meant as a partition of the set of solutions having certain properties listed below. Each cluster contains an exponential number of solutions,  $\exp[Ns_{\text{int}}]$ , and the clusters are themselves exponentially numerous,  $\exp[N\Sigma]$ . The total entropy density thus decomposes into the sum of  $s_{\text{int}}$ , the internal entropy of the clusters and  $\Sigma$ , encoding the degeneracy of these clusters, usually termed complexity in this context. Furthermore, solutions inside a given cluster should be well-connected, while two

solutions of distinct clusters are well-separated. A possible definition for these notions is the following. Suppose  $\underline{\sigma}$  and  $\underline{\tau}$  are two solutions of a given cluster. Then one can construct a path  $(\underline{\sigma} = \underline{\sigma}_0, \underline{\sigma}_1, \dots, \underline{\sigma}_{n-1}, \underline{\sigma}_n = \underline{\tau})$  where any two successive  $\underline{\sigma}_i$  are separated by a sub-extensive Hamming distance. On the contrary such a path does not exist if  $\underline{\sigma}$  and  $\underline{\tau}$  belong to two distinct clusters. Clustered configuration spaces as described above have been often encountered in various contexts, e.g. neural networks [23] and mean-field spin glasses [24]. A vast body of involved, yet non-rigorous, analytical techniques [1] have been developed in the field of statistical mechanics of disordered systems to tackle such situations, some of them having been justified rigorously [25–27]. In this literature clusters appear under the name of “pure states”, or “lumps” (see for instance the chapter 6 of [25] for a rigorous definition and proof of existence in a related model). As we shall explain in a few lines, this clustering phenomenon has been demonstrated rigorously in the case of random XORSAT instances [28, 29]. For random SAT instances, where in fact the detailed picture of the satisfiable phase is thought to be richer [22], there are some rigorous results [30–32] on the existence of clusters for large enough  $k$ .

## B. Phase transitions in random XORSAT

Consider an instance  $F$  of the XORSAT problem [33], i.e. a list of  $M$  linear equations each involving  $k$  out of  $N$  boolean variables, where the additions are computed modulo 2. The study performed in [28, 29] provides a detailed picture of the clustering and satisfiability transition sketched above. A crucial point is the construction of a core subformula according to the following algorithm. Let us denote  $F_0 = F$  the initial set of equations, and  $V_0$  the set of variables which appear in at least one equation of  $F_0$ . A sequence  $F_T, V_T$  is constructed recursively: if there are no variables in  $V_T$  which appear in exactly one equation of  $F_T$  the algorithm stops. Otherwise one of these “leaf variables”  $\sigma_i$  is chosen arbitrarily,  $F_{T+1}$  is constructed from  $F_T$  by removing the unique equation in which  $\sigma_i$  appeared, and  $V_{T+1}$  is defined as the set of variables which appear at least once in  $F_{T+1}$ . Let us call  $T_*$  the number of steps performed before the algorithm stops, and  $F' = F_{T_*}$ ,  $V' = V_{T_*}$  the remaining clauses and variables. Note first that despite the arbitrariness in the choice of the removed leaves, the output subformula  $F'$  is unambiguously determined by  $F$ . Indeed,  $F'$  can be defined as the maximal (in the inclusion sense) subformula in which all present variables have a minimal occurrence number of 2, and is thus unique. In graph theoretic terminology  $F'$  is the 2-core of  $F$ , the  $q$ -core of hypergraphs being a generalization of the more familiar notion on graphs, thoroughly studied in random graph ensembles in [34]. Extending this study, relying on the approximability of this leaf removal process by differential equations [35], it was shown in [28, 29] that there is a threshold phenomenon at  $\alpha_d(k)$ . For  $\alpha < \alpha_d$  the 2-core  $F'$  is, with high probability, empty, whereas it contains a finite fraction of the variables and equations for  $\alpha > \alpha_d$ .  $\alpha_d$  is easily determined numerically: it is the smallest value of  $\alpha$  such that the equation  $x = 1 - \exp[-\alpha k x^{k-1}]$  has a non-trivial solution in  $(0, 1]$ .

It turns out that  $F$  is satisfiable if and only if  $F'$  is, and that the number of solutions of these two formulas are related in an enlightening way. It is clear that if the 2-core has no solution, there is no way to find one for the full formula. Suppose on the contrary that an assignment of the variables in  $V'$  that satisfy the equations of  $F'$  has been found, and let us show how to construct a solution of  $F$  (and count in how many possible ways we can do this). Set  $\mathcal{N}_0 = 1$ , and reintroduce step by step the removed equations, starting from the last: in the  $n$ 'th step of this new procedure we reintroduce the clause which was removed at step  $T_* - n$  of the leaf removal. This reintroduced clause has  $d_n = |V_{T_*-n-1}| - |V_{T_*-n}| \geq 1$  leaves; their configuration can be chosen in  $2^{d_n-1}$  ways to satisfy the reintroduced clause, irrespectively of the previous choices, and we bookkeep this number of possible extensions by setting  $\mathcal{N}_{n+1} = \mathcal{N}_n 2^{d_n-1}$ . Finally the total number of solutions of  $F$  compatible with the choice of the solution of  $F'$  is obtained by adding the freedom of the variables which appeared in no equations of  $F$ ,  $\mathcal{N}_{\text{int}} = \mathcal{N}_{T_*} 2^{N-|V_0|}$ . Let us underline that  $\mathcal{N}_{\text{int}}$  is independent of the initial satisfying assignment of the variables in  $V'$ , as appears clearly from the description of the reconstruction algorithm; this property can be traced back to the linear algebra structure of the problem. This suggests naturally the decomposition of the total number of solutions of  $F$  as the product of the number of satisfying assignments of  $V'$ , call it  $\mathcal{N}_{\text{core}}$ , by the number of compatible full solutions  $\mathcal{N}_{\text{int}}$ . In terms of the associated entropy densities this decomposition is additive

$$s = \Sigma + s_{\text{int}} , \quad \Sigma \equiv \frac{1}{N} \ln \mathcal{N}_{\text{core}} , \quad s_{\text{int}} \equiv \frac{1}{N} \ln \mathcal{N}_{\text{int}} , \quad (22)$$

where the quantity  $\Sigma$  is the entropy density associated to the core of the formula. It is in fact much easier technically to compute the statistical (with respect to the choice of the random formula  $F$ ) properties of  $\Sigma$  and  $s_{\text{int}}$  once this decomposition has been done (the fluctuations in the number of solutions is much smaller once the non-core part of the formula has been removed). The outcome of the computations [28, 29] is the determination of the threshold value  $\alpha_s$  for the appearance of a solution of the 2-core  $F'$  (and thus of the complete formula), along with explicit formulas for the typical values of  $\Sigma$  and  $s$ . These two quantities are plotted on Fig. 2. The satisfiability threshold corresponds



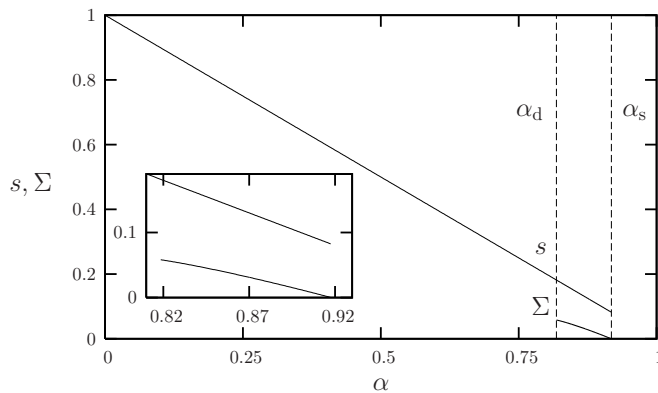


FIG. 2: Complexity and total entropy for 3-XORSAT, in units of  $\ln 2$ . The inset presents an enlargement of the regime  $\alpha \in [\alpha_d, \alpha_s]$ .

TABLE I: Critical connectivities for the dynamical, condensation and satisfiability transitions for  $k$ -SAT random formulas.

	$\alpha_d$ [22]	$\alpha_c$ [22]	$\alpha_s$ [38]
$k = 3$	3.86	3.86	4.267
$k = 4$	9.38	9.547	9.93
$k = 5$	19.16	20.80	21.12
$k = 6$	36.53	43.08	43.4

to the cancellation of  $\Sigma$ : the number of solutions of the core vanishes continuously at  $\alpha_s$ , while the total entropy remains finite because of the freedom of choice for the variables in the non-core part of the formula.

On top of the simplification in the analytical determination of the satisfiability threshold, this core decomposition of a formula unveils the change in the structure of the set of solutions that occurs at  $\alpha_d$ . Indeed, let us call cluster all solutions of  $F$  reconstructed from a common solution of  $F'$ . Then one can show that this partition of the solution set of  $F$  exhibits the properties exposed in Sec. III A, namely that solutions are well-connected inside a cluster and separated from one cluster to another. The number of clusters is precisely equal to the number of solutions of the core subformula, it thus undergoes a drastic modification at  $\alpha_d$ . For smaller ratio of constraints the core is typically empty, there is one single cluster containing all solutions; when the threshold  $\alpha_d$  is reached there appears an exponential numbers of clusters, the rate of growth of this exponential being given by the complexity  $\Sigma$ . Before considering the extension of this picture to random SAT problems, let us mention that further studies of the geometry of the space of solutions of random XORSAT instances can be found in [36, 37].

### C. Phase transitions in random SAT

The possibility of a clustering transition in random SAT problems was first studied in [20] by means of variational approximations. Later developments allowed the computation of the complexity and, from the condition of its cancellation, the estimation of the satisfiability threshold  $\alpha_s$ . This was first done for  $k = 3$  in [21] and generalized for  $k \geq 4$  in [38], some of the values of  $\alpha_s$  thus computed are reported in Tab. I. A systematic expansion of  $\alpha_s$  at large  $k$  was also performed in [38].

SAT formulas do not share the linear algebra structure of XORSAT, which makes the analysis of the clustering transition much more difficult, and leads to a richer structure of the satisfiable phase  $\alpha \leq \alpha_s$ . The simple graph theoretic arguments are not valid anymore, one cannot extract a core subformula from which the partition of the solutions into clusters follows directly. It is thus necessary to define them as a partition of the solutions such that each cluster is well-connected and well-separated from the other ones. A second complication arises: there is no reason for the clusters to contain all the same number of solutions, as was ensured by the linear structure of XORSAT. On the contrary, as was observed in [20] and in [39] for the similar random COL problem, one faces a variety of clusters with various internal entropies  $s_{\text{int}}$ . The complexity  $\Sigma$  becomes a function of  $s_{\text{int}}$ , in other words the number of clusters of internal entropy density  $s_{\text{int}}$  is typically exponential, growing at the leading order like  $\exp[N\Sigma(s_{\text{int}})]$ . Drawing the consequences of these observations, a refined picture of the satisfiable phase, and in particular the existence of a

new (so-called condensation) threshold  $\alpha_c \in [\alpha_d, \alpha_s]$ , was advocated in [22]. Let us briefly sketch some of these new features and their relationship with the previous results of [21, 38]. Assuming the existence of a positive, concave, complexity function  $\Sigma(s_{\text{int}})$ , continuously vanishing outside an interval of internal entropy densities  $[s_-, s_+]$ , the total entropy density is given by

$$s = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \int_{s_-}^{s_+} ds_{\text{int}} e^{N[\Sigma(s_{\text{int}}) + s_{\text{int}}]} . \quad (23)$$

In the thermodynamic limit the integral can be evaluated with the Laplace method. Two qualitatively distinct situations can arise, whether the integral is dominated by a critical point in the interior of the interval  $[s_-, s_+]$ , or by the neighborhood of the upper limit  $s_+$ . In the former case an overwhelming majority of the solutions are contained in an exponential number of clusters, while in the latter the dominant contributions comes from a sub-exponential number of clusters of internal entropy  $s_+$ , as  $\Sigma(s_+) = 0$ . The threshold  $\alpha_c$  separates the first regime  $[\alpha_d, \alpha_c]$  where the relevant clusters are exponentially numerous, from the second, condensed situation for  $\alpha \in [\alpha_c, \alpha_s]$  with a sub-exponential number of dominant clusters<sup>3</sup>.

The computations of [21, 38] did not take into account the distribution of the various internal entropies of the clusters, which explains the discrepancy in the estimation of the clustering threshold  $\alpha_d$  between [21, 38] and [22]. Let us however emphasize that this refinement of the picture does not contradict the estimation of the satisfiability threshold of [21, 38]: the complexity computed in these works is  $\Sigma_{\text{max}}$ , the maximal value of  $\Sigma(s_{\text{int}})$  reached at a local maximum with  $\Sigma'(s) = 0$ , which indeed vanishes when the whole complexity function disappears.

It is fair to say that the details of the picture proposed by statistical mechanics studies have rapidly evolved in the last years, and might still be improved. They rely indeed on self-consistent assumptions which are rather tedious to check [40]. Some elements of the clustering scenario have however been established rigorously in [30–32], at least for large enough  $k$ . In particular these works demonstrated, for some values of  $k$  and  $\alpha$  in the satisfiable regime, the existence of forbidden intermediate Hamming distances between pairs of configurations, which are either close (in the same cluster) or far apart (in two distinct clusters).

Note finally that the consequences of such distributions of clusters internal entropies were investigated on a toy model in [41], and that yet another threshold  $\alpha_f > \alpha_d$  for the appearance of frozen variables constrained to take the same values in all solutions of a given cluster was investigated in [42].

#### D. A glimpse at the computations

The statistical mechanics of disordered systems [1] was first developed on so-called fully-connected models, where each variable appears in a number of constraints which diverges in the thermodynamic limit. This is for instance the case of the perceptron problem discussed in Sec. II. On the contrary, in a random  $k$ -SAT instance a variable is typically involved in a finite number of clauses, one speaks in this case of a diluted model. This finite connectivity is a source of major technical complications. In particular the replica method, alluded to in Sec. IIC and applied to random  $k$ -SAT in [19, 20], turns out to be rather cumbersome for diluted models in the presence of clustering [43]. The cavity formalism [21, 44, 45], formally equivalent to the replica one, is more adapted to the diluted models. In the following paragraphs we shall try to give a few hints at the strategy underlying the cavity computations, that might hopefully ease the reading of the original literature.

The description of the random formula ensemble has two complementary aspects: a global (thermodynamic) one, which amounts to the computation of the typical energy and number of optimal configurations. A more ambitious description will also provide geometrical information on the organization of this set of optimal configurations inside the  $N$ -dimensional hypercube. As discussed above these two aspects are in fact interleaved, the clustering affecting both the thermodynamics (by the decomposition of the entropy into the complexity and the internal entropy) and the geometry of the configuration space. Let us for simplicity concentrate on the  $\alpha < \alpha_s$  regime and consider a satisfiable formula  $F$ . Both thermodynamic and geometric aspects can be studied in terms of the uniform probability law over the solutions of  $F$ :

$$\mu(\underline{\sigma}) = \frac{1}{Z} \prod_{a=1}^M w_a(\underline{\sigma}_a) , \quad (24)$$

---

<sup>3</sup> This picture is expected to hold for  $k \geq 4$ ; for  $k = 3$ , the dominant clusters are expected to be of sub-exponential degeneracy in the whole clustered phase, hence  $\alpha_c = \alpha_d$  in this case.

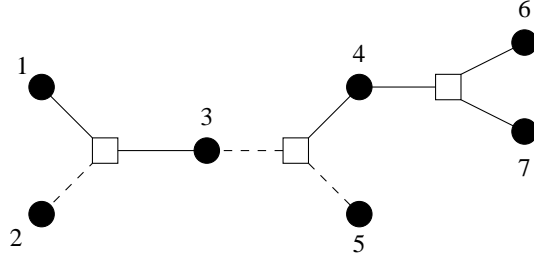


FIG. 3: The factor graph representation of a small 3-SAT formula:  $(x_1 \vee \overline{x_2} \vee x_3) \wedge (\overline{x_3} \vee x_4 \vee \overline{x_5}) \wedge (x_4 \vee x_6 \vee x_7)$ .

where  $Z$  is the number of solutions of  $F$ , the product runs over its clauses, and  $w_a$  is the indicator function of the event “clause  $a$  is satisfied by the assignment  $\underline{\sigma}$ ” (in fact this depends only on the configuration of the  $k$  variables involved in the clause  $a$ , that we denote  $\underline{\sigma}_a$ ). For instance the (information theoretic) entropy of  $\mu$  is equal to  $\ln Z$ , the log degeneracy of solutions, and geometric properties can be studied by computing averages with respect to  $\mu$  of well-chosen functions of  $\underline{\sigma}$ .

A convenient representation of such a law is provided by factor graphs [46]. These are bipartite graphs with two types of vertices (see Fig. 3 for an illustration): one variable node (filled circle) is associated to each of the  $N$  Boolean variables, while the clauses are represented by  $M$  constraint nodes (empty squares). By convention we use the indices  $a, b, \dots$  for the constraint nodes,  $i, j, \dots$  for the variables. An edge is drawn between variable node  $i$  and constraint node  $a$  if and only if  $a$  depends on  $i$ . To precise further by which value of  $\sigma_i$  the clause  $a$  gets satisfied one can use two type of linestyles, solid and dashed on the figure. A notation repeatedly used in the following is  $\partial a$  (resp.  $\partial i$ ) for the neighborhood of a constraint (resp. variable) node, i.e. the set of adjacent variable (resp. constraint) nodes. In this context  $\setminus$  denotes the subtraction from a set. We shall more precisely denote  $\partial_+ i(a)$  (resp.  $\partial_- i(a)$ ) the set of clauses in  $\partial i \setminus a$  agreeing (resp. disagreeing) with  $a$  on the satisfying value of  $\sigma_i$ , and  $\partial_\sigma i$  the set of clauses in  $\partial i$  which are satisfied by  $\sigma_i = \sigma$ . This graphical representation naturally suggests a notion of distance between variable nodes  $i$  and  $j$ , defined as the minimal number of constraint nodes crossed on a path of the factor graph linking nodes  $i$  and  $j$ .

Suppose now that  $F$  is drawn from the random ensemble. The corresponding random factor graph enjoys several interesting properties [7]. The degree  $|\partial i|$  of a randomly chosen variable  $i$  is, in the thermodynamic limit, a Poisson random variable of average  $\alpha k$ . If instead of a node one chooses randomly an edge  $a - i$ , the outdegree  $|\partial i \setminus a|$  of  $i$  has again a Poisson distribution with the same parameter. Moreover the sign of the literals being chosen uniformly, independently of the topology of the factor graph, the degrees  $|\partial_+ i|$ ,  $|\partial_- i|$ ,  $|\partial_+ i(a)|$  and  $|\partial_- i(a)|$  are Poisson random variables of parameter  $\alpha k/2$ . Another important feature of these random factor graphs is their local tree-like character: if the portion of the formula at graph distance smaller than  $L$  of a randomly chosen variable is exposed, the probability that this subgraph is a tree goes to 1 if  $L$  is kept fixed while the size  $N$  goes to infinity.

Let us for a second forget about the rest of the graph and consider a finite formula whose factor graph is a tree, as is the case for the example of Fig. 3. The probability law  $\mu$  of Eq. (24) becomes in this case a rather simple object. Tree structures are indeed naturally amenable to a recursive (dynamic programming) treatment, operating first on sub-trees which are then glued together. More precisely, for each edge between a variable node  $i$  and a constraint node  $a$  one defines the amputated tree  $F_{a \rightarrow i}$  (resp.  $F_{i \rightarrow a}$ ) by removing all clauses in  $\partial i$  apart from  $a$  (resp. removing only  $a$ ). These subtrees are associated to probability laws  $\mu_{a \rightarrow i}$  (resp.  $\mu_{i \rightarrow a}$ ), defined as in Eq. (24) but with a product running only on the clauses present in  $F_{a \rightarrow i}$  (resp.  $F_{i \rightarrow a}$ ). The marginal law of the root variable  $i$  in these amputated probability measures can be parametrized by a single real, as  $\sigma_i$  can take only two values (that, in the Ising spin convention, are  $\pm 1$ ). We thus define these fields, or messages,  $h_{i \rightarrow a}$  and  $u_{a \rightarrow i}$ , by

$$\mu_{i \rightarrow a}(\sigma_i) = \frac{1 - J_i^a \sigma_i \tanh h_{i \rightarrow a}}{2}, \quad \mu_{a \rightarrow i}(\sigma_i) = \frac{1 - J_i^a \sigma_i \tanh u_{a \rightarrow i}}{2}, \quad (25)$$

where we recall that  $\sigma_i = J_i^a$  is the value of the literal  $i$  unsatisfying clause  $a$ . A standard reasoning (see for instance [47]) allows to derive recursive equations (illustrated in Fig. 4) on these messages,

$$h_{i \rightarrow a} = \sum_{b \in \partial_+ i(a)} u_{b \rightarrow i} - \sum_{b \in \partial_- i(a)} u_{b \rightarrow i}, \quad (26)$$

$$u_{a \rightarrow i} = -\frac{1}{2} \ln \left( 1 - \prod_{j \in \partial a \setminus i} \frac{1 - \tanh h_{j \rightarrow a}}{2} \right).$$

Because the factor graph is a tree this set of equations has a unique solution which can be efficiently determined: one start from the leaves (degree 1 variable nodes) which obey the boundary condition  $h_{i \rightarrow a} = 0$ , and progresses inwards the graph. The law  $\mu$  can be completely described from the values of the  $h$ 's and  $u$ 's solutions of these equations for all edges of the graph. For instance the marginal probability of  $\sigma_i$  can be written as

$$\mu(\sigma_i) = \frac{1 + \sigma_i \tanh h_i}{2}, \quad h_i = \sum_{a \in \partial_+ i} u_{a \rightarrow i} - \sum_{a \in \partial_- i} u_{a \rightarrow i}. \quad (27)$$

In addition the entropy  $s$  of solutions of such a tree formula, can be computed from the values of the messages  $h$  and  $u$  [47].

We shall come back to the equations (26), and justify the denomination messages, in Sec. VC; these can be interpreted as the Belief Propagation [46, 48, 49] heuristic equations for loopy factor graphs.

The factor graph of random formulas is only locally tree-like; the simple computation sketched above has thus to be amended in order to take into account the effect of the distant, loopy part of the formula. Let us call  $F_L$  the factor graph made of variable nodes at graph distance smaller than or equal to  $L$  from an arbitrarily chosen variable node  $i$  in a large random formula  $F$ , and  $B_L$  the variable nodes at distance exactly  $L$  from  $i$ . Without loss of generality in the thermodynamic limit, we can assume that  $F_L$  is a tree. The cavity method amounts to an hypothesis on the effect of the distant part of the factor graph,  $F \setminus F_L$ , i.e. on the boundary condition it induces on  $F_L$ . In its simplest (so called replica symmetric) version, that is believed to correctly describe the unclustered situation for  $\alpha \leq \alpha_d$ ,  $F \setminus F_L$  is replaced, for each variable node  $j$  in the boundary  $B_L$ , by a fictitious constraint node which sends a bias  $u_{\text{ext} \rightarrow j}$ . In other words the boundary condition is factorized on the various nodes of  $B_L$ ; such a simple description is expected to be correct for  $\alpha \leq \alpha_d$  because, in the amputated factor graph  $F \setminus F_L$ , the distance between the variables of  $B_L$  is typically large (of order  $\ln N$ ), and these variables should thus be weakly correlated. These external biases are then turned into random variables to take into account the randomness in the construction of the factor graphs, and Eq. (26) acquires a distributional meaning. The messages  $h$  (resp.  $u$ ) are supposed to be i.i.d. random variables drawn from a common distribution, the degrees  $\partial_{\pm} i(a)$  being two independent Poisson random variables of parameter  $\alpha k/2$ . These distributional equations can be numerically solved by a population dynamics algorithm [44], also known as a particle representation in the statistics litterature. The typical entropy density is then computed by averaging  $s$  over these distributions of  $h$  and  $u$ .

This description fails in the presence of clustering, which induces correlations between the variable nodes of  $B_L$  in the amputated factor graph  $F \setminus F_L$ . To take these correlations into account a refined version of the cavity method (termed one step of replica symmetry breaking, in short 1RSB) has been developed. It relies on the hypothesis that the partition of the solution space into clusters  $\gamma$  has nice decorrelation properties: once decomposed onto this partition,  $\mu$  restricted to a cluster  $\gamma$  behaves essentially as in the unclustered phase (it is a pure state in statistical mechanics jargon). Each directed edge  $a \rightarrow i$  should thus bear a family of messages  $u_{a \rightarrow i}^\gamma$ , one for each cluster, or alternatively a distribution  $Q_{a \rightarrow i}(u)$  of the messages with respect to the choice of  $\gamma$ . The equations (26) are thus promoted to recursions between distributions  $P_{i \rightarrow a}(h)$ ,  $Q_{a \rightarrow i}(u)$ , which depends on a real  $m$  known as the Parisi breaking parameter. Its role is to select the size of the investigated clusters, i.e. the number of solutions they contain. The computation of the typical entropy density is indeed replaced by a more detailed thermodynamic potential,

$$\Phi(m) = \frac{1}{N} \ln \sum_{\gamma} Z_{\gamma}^m = \frac{1}{N} \ln \int_{s_-}^{s_+} ds_{\text{int}} e^{N[\Sigma(s_{\text{int}}) + m s_{\text{int}}]}. \quad (28)$$

In this formula  $Z_{\gamma}$  denotes the number of solutions inside a cluster  $\gamma$ , and we used the hypothesis that at the leading order the number of clusters with internal entropy density  $s_{\text{int}}$  is given by  $\exp[N\Sigma(s_{\text{int}})]$ . The complexity function  $\Sigma(s_{\text{int}})$  can thus be obtained from  $\Phi(m)$  by an inverse Legendre transform. For generic values of  $m$  this approach is computationally very demanding; following the same steps as in the replica symmetric version of the cavity method one faces a distribution (with respect to the topology of the factor graph) of distributions (with respect to the choice of the clusters) of messages. Simplifications however arise for  $m = 1$  and  $m = 0$  [22]; the latter case corresponds in fact to the original Survey Propagation approach of [21]. As appears clearly in Eq. (28), for this value of  $m$  all clusters are treated on an equal footing and the dominant contribution comes from the most numerous clusters, independently of their sizes. Moreover, as we further explain in Sec. VC, the structure of the equations can be greatly simplified in this case, the distribution over the cluster of fields being parametrized by a single number.

## E. Finite Size Scaling results

As we explained in Sec. IIB the threshold phenomenon can be more precisely described by finite size scaling relations. Let us mention some FSS results about the transitions we just discussed.

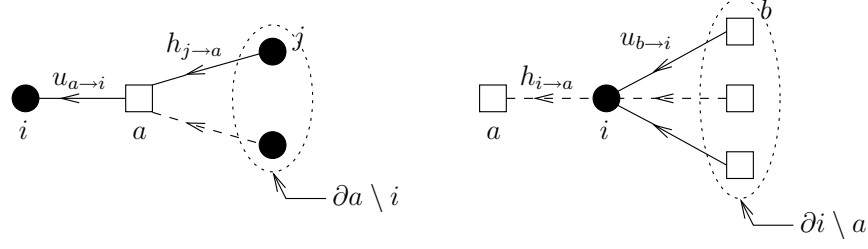


FIG. 4: A schematic representation of Eq. (26).

For random 2-SAT, where the satisfiability property is known [50] to exhibit a sharp threshold at  $\alpha_s = 1$ , the width of the transition window has been determined in [51]. The range of  $\alpha$  where the probability of satisfaction drops significantly is of order  $N^{-1/3}$ , i.e. the exponent  $\nu$  is equal to 3, as for the random graph percolation. This similarity is not surprising, the proof of [51] relies indeed on a mapping of 2-SAT formulas onto random (directed) graphs.

The clustering transition for XORSAT was first conjectured in [52] (in the related context of error-correcting codes) then proved in [53] to be described by

$$P(N, M = N(\alpha_d + N^{-1/2}\lambda + N^{-2/3}\delta)) = \mathcal{F}(\lambda) + O(N^{-5/26}), \quad (29)$$

where  $\delta$  is a subleading shift correction that has been explicitly computed, and the scaling function  $\mathcal{F}$  is, upto a multiplicative factor on  $\lambda$ , the same error function as in Eq. (4).

A general result has been proved in [54] on the width of transition windows. Under rather unrestrictive conditions one can show that  $\nu \geq 2$ : the transitions cannot be arbitrarily sharp. Roughly speaking the bound is valid when a finite fraction of the clauses are not decisive for the property of the formulas studied, for instance clauses containing a leaf variable are not relevant for the satisfiability of a formula. The number of these irrelevant clauses is of order  $N$  and has thus natural fluctuations of order  $\sqrt{N}$ ; these fluctuations blur the transition window which cannot be sharper than  $N^{-1/2}$ .

Several studies (see for instance [33, 55, 56]) have attempted to determine the transition window from numeric evaluations of the probability  $P(N, \alpha)$ , for instance for the satisfiability threshold of random 3-SAT [55, 56] and XORSAT [33]. These studies are necessarily confined to small formula sizes, as the typical computation cost of complete algorithms grows exponentially around the transition. In consequence the asymptotic regime of the transition window,  $N^{-1/\nu}$ , is often hidden by subleading corrections which are difficult to evaluate, and in [55, 56] the reported values of  $\nu$  were found to be in contradiction with the latter derived rigorous bound. This is not an isolated case, numerical studies are often plagued by uncontrolled finite-size effects, as for instance in the bootstrap percolation [57], a variation of the classical percolation problem.

#### IV. LOCAL SEARCH ALGORITHMS

The following of this review will be devoted to the study of various solving algorithms for SAT formulas. Algorithms are, to some extent, similar to dynamical processes studied in statistical physics. In this context the focus is however mainly on stochastic processes that respect detailed balance with respect to the Gibbs-Boltzmann measure [58], a condition which is rarely respected by solving algorithms. Physics inspired techniques can yet be useful, and will emerge in three different ways. The random walk algorithms considered in this Section are stochastic processes in the space of configurations (not fulfilling the detailed balance condition), moving by small steps where one or a few variables are modified. Out-of-equilibrium physics (and in particular growth processes) provide an interesting view on classical complete algorithms (DPLL), as shown in Sec. VB. Finally, the picture of the satisfiable phase put forward in Sec. III underlies the message-passing procedures discussed in Sec. VC.

##### A. Pure random walk sat, definition and results valid for all instances

Papadimitriou [59] proposed the following algorithm, called Pure Random Walk Sat (PRWSAT) in the following, to solve  $k$ -SAT formulas:

1. Choose an initial assignment  $\underline{\sigma}(0)$  uniformly at random and set  $T = 0$ .

2. If  $\underline{\sigma}(T)$  is a solution of the formula (i.e.  $E(\underline{\sigma}(T)) = 0$ ), output SOLUTION and stop. If  $T = T_{\max}$ , a threshold fixed beforehand, output UNDETERMINED and stop.
3. Otherwise, pick uniformly at random a clause among those that are UNSAT in  $\underline{\sigma}(T)$ ; pick uniformly at random one of the  $k$  variables of this clause and flip it (reverse its status from True to False and vice-versa) to define the next assignment  $\underline{\sigma}(T+1)$ ; set  $T \rightarrow T+1$  and go back to step 2.

This defines a stochastic process  $\underline{\sigma}(T)$ , a biased random walk in the space of configurations. The modification  $\underline{\sigma}(T) \rightarrow \underline{\sigma}(T+1)$  in step 3 makes the selected clause satisfied; however the flip of a variable  $i$  can turn previously satisfied clauses into unsatisfied ones (those which were satisfied solely by  $i$  in  $\underline{\sigma}(T)$ ).

This algorithm is not complete: if it outputs a solution one is certain that the formula was satisfiable (and the current configuration provides a certificate of it), but if no solution has been found within the  $T_{\max}$  allowed steps one cannot be sure that the formula was unsatisfiable. There are however two rigorous results which makes it a probabilistically almost complete algorithm [60].

For  $k = 2$ , it was shown in [59] that PRWSAT finds a solution in a time of order  $O(N^2)$  with high probability for all satisfiable instances. Hence, one is almost certain that the formula was unsatisfiable if the output of the algorithm is UNDETERMINED after  $T_{\max} = O(N^2)$  steps.

Schöning [61] proposed the following variation for  $k = 3$ . If the algorithm fails to find a solution before  $T_{\max} = 3N$  steps, instead of stopping and printing UNDETERMINED, it restarts from step 1, with a new random initial condition  $\underline{\sigma}(0)$ . Schöning proved that if after  $R$  restarts no solution has been found, then the probability that the instance is satisfiable is upper-bounded by  $\exp[-R \times (3/4)^N]$  (asymptotically in  $N$ ). This means that a computational cost of order  $(4/3)^N$  allows to reduce the probability of error of the algorithm to arbitrary small values. Note that if the time scaling of this bound is exponential, it is also exponentially smaller than the  $2^N$  cost of an exhaustive enumeration. Improvements on the factor  $4/3$  are reported in [62].

## B. Typical behavior on random $k$ -SAT instances

The results quoted above are true for any  $k$ -SAT instance. An interesting phenomenology arises when one applies the PRWSAT algorithm to instances drawn from the random  $k$ -SAT ensemble [63, 64]. Figure 5 displays the temporal evolution of the number of unsatisfied clauses during the execution of the algorithm, for two random 3-SAT instances of constraint ratio  $\alpha = 2$  and 3. The two curves are very different: at low values of  $\alpha$  the energy decays rather fast towards 0, until a point where the algorithm finds a solution and stops. On the other hand, for larger values of  $\alpha$ , the energy first decays towards a strictly positive value, around which it fluctuates for a long time, until a large fluctuation reaches 0, signaling the discovery of a solution. A more detailed study with formulas of increasing sizes reveals that a threshold value  $\alpha_{\text{rw}} \approx 2.7$  (for  $k = 3$ ) sharply separates this two dynamical regimes. In fact the fraction of unsatisfied clauses  $\varphi = E/M$ , expressed in terms of the reduced time  $t = T/M$ , concentrates in the thermodynamic limit around a deterministic function  $\varphi(t)$ . For  $\alpha < \alpha_{\text{rw}}$  the function  $\varphi(t)$  reaches 0 at a finite value  $t_{\text{sol}}(\alpha, k)$ , which means that the algorithm finds a solution in a linear number of steps, typically close to  $Nt_{\text{sol}}(\alpha, k)$ . On the contrary for  $\alpha > \alpha_{\text{rw}}$  the reduced energy  $\varphi(t)$  reaches a positive value  $\varphi_{\text{as}}(\alpha, k)$  as  $t \rightarrow \infty$ ; a solution, if any, can be found only through large fluctuations of the energy which occur on a time scale exponentially large in  $N$ . This is an example of a metastability phenomenon, found in several other stochastic processes, for instance the contact process [65]. When the threshold  $\alpha_{\text{rw}}$  is reached from below the solving time  $t_{\text{sol}}(\alpha, k)$  diverges, while the height of the plateau  $\varphi_{\text{as}}(\alpha, k)$  vanishes when  $\alpha_{\text{rw}}$  is approached from above.

In [63, 64] various statistical mechanics inspired techniques have been applied to study analytically this phenomenology, some results are presented in Figure 6. The low  $\alpha$  regime can be tackled by a systematic expansion of  $t_{\text{sol}}(\alpha, k)$  in powers of  $\alpha$ . The first three terms of these series have been computed, and are shown on the left panel to be in good agreement with the numerical simulations.

Another approach was followed to characterize the transition  $\alpha_{\text{rw}}$ , and to compute (approximations of) the asymptotic fraction of unsatisfied clauses  $\varphi_{\text{as}}$  and the intensity of the fluctuations around it. The idea is to project the Markovian evolution of the configuration  $\underline{\sigma}(T)$  on a simpler observable, the energy  $E(T)$ . Obviously the Markovian property is lost in this transformation, and the dynamics of  $E(T)$  is much more complex. One can however approximate it by assuming that all configurations of the same energy  $E(T)$  are equiprobable at a given step of execution of the algorithm. This rough approximation of the evolution of  $E(T)$  is found to concentrate around its mean value in the thermodynamic limit, as was constated numerically for the original process. Standard techniques allow to compute this average approximated evolution, which exhibits the threshold behavior explained above at a value  $\alpha = (2^k - 1)/k$  which is, for  $k = 3$ , slightly lower than the threshold  $\alpha_{\text{rw}}$ . The right panel of Fig. 6 confronts the results of this approximation with the numerical simulations; given the roughness of the hypothesis the agreement is rather satisfying, and is expected to improve for larger values of  $k$ .

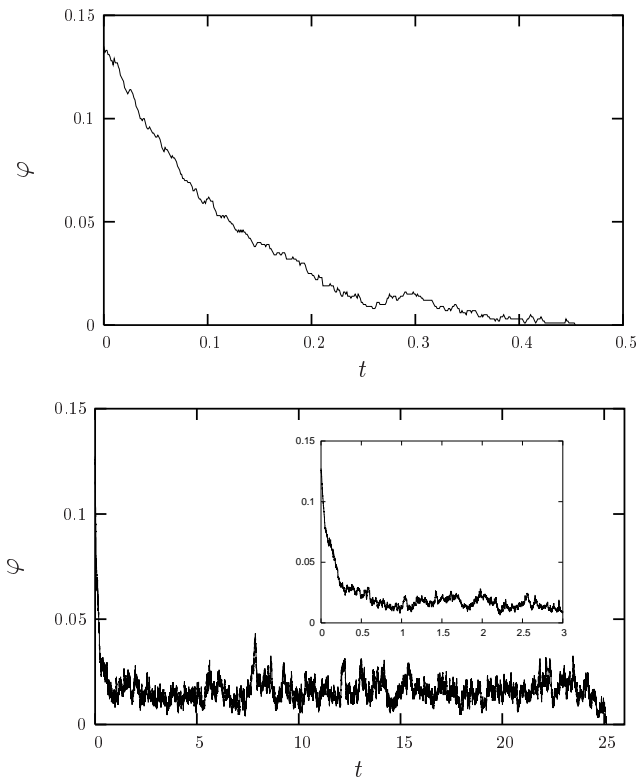


FIG. 5: Fraction of unsatisfied constraints  $\varphi = E/M$  in function of reduced time  $t = T/M$  during the execution of PRWSAT on random 3-SAT formulas with  $N = 500$  variables. Top:  $\alpha = 2$ , Bottom:  $\alpha = 3$ .

The rigorous results on the behavior of PRWSAT on random instances are very few. Let us mention in particular [66], which proved that the solving time for random 3-SAT formulas is typically polynomial up to  $\alpha = 1.63$ , a result in agreement yet weaker than the numerical results presented here.

### C. More performant variants of the algorithm

The threshold  $\alpha_{\text{rw}}$  for linear time solving of random instances by PRWSAT was found above to be much smaller than the satisfiability threshold  $\alpha_s$ . It must however be emphasized that PRWSAT is only the simplest example of a large family of local search algorithms, see for instance [67–71]. They all share the same structure: a solution is searched through a random walk in the space of configurations, one variable being modified at each step. The choice of the flipped variable is made according to various heuristics; the goal is to find a compromise between the greediness of the walk which seeks to minimize locally the energy of the current assignment, and the necessity to allow for moves increasing the energy in order to avoid the trapping in local minima of the energy function. A frequently encountered ingredient of the heuristics, which is of a greedy nature, is the focusing: the flipped variable necessarily belongs to at least one unsatisfied clause before the flip, which thus becomes satisfied after the move. Moreover, instead of choosing randomly one of the  $k$  variables of the unsatisfied clause, one can consider for each of them the effect of the flip, and avoid variables which, once flipped, will turn satisfied clauses into unsatisfied ones [67, 68]. Another way to implement the greediness [69] consists in bookkeeping the lowest energy found so far during the walk, and forbids flips which will raise the energy of the current assignment above the registered record plus a tolerance threshold. These demanding requirements have to be balanced with noisy, random steps, allowing to escape traps which are only locally minima of the objective function.

These more elaborated heuristics are very numerous, and depend on parameters that are finely tuned to achieve the best performances, hence an exhaustive comparison is out of the scope of this review. Let us only mention that some of these heuristics are reported in [69, 70] to efficiently find solutions of large (up to  $N = 10^6$ ) random formulas of 3-SAT at ratio  $\alpha$  very close to the satisfiability threshold, i.e. for  $\alpha \lesssim 4.21$ .

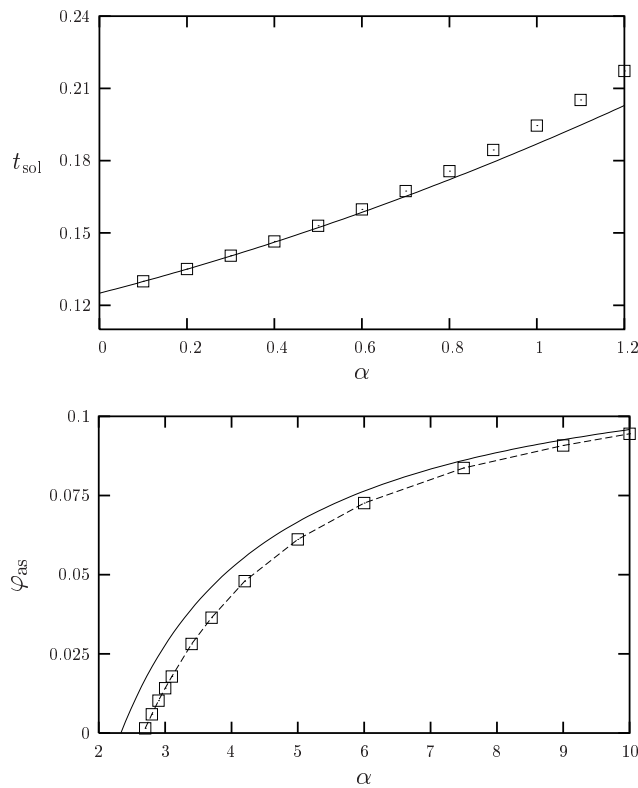


FIG. 6: Top: linear solving time  $t_{\text{sol}}(\alpha, 3)$  for random 3-SAT formulas in function of  $\alpha$ ; symbols correspond to numerical simulations, solid line to the second order expansion in  $\alpha$  obtained in [63]. Bottom: fraction of unsatisfied constraints reached at large time for  $\alpha > \alpha_{\text{rw}}$  for random 3-SAT formulas; symbols correspond to numerical simulations, solid line to the approximate analytical computations of [63, 64].

## V. DECIMATION BASED ALGORITHMS

The algorithms studied in the remaining of the review are of a very different nature compared to the local search procedures described above. Given an initial formula  $F$  whose satisfiability has to be decided, they proceed by assigning sequentially the value of some of the variables. The formula can be simplified under such a partial assignment: clauses which are satisfied by at least one of their literal can be removed, while literals unsatisfying a clause are discarded from the clause. It is instructive to consider the following thought experiment: suppose one can consult an oracle who, given a formula, is able to compute the marginal probability of the variables, in the uniform probability measure over the optimal assignments of the formula. With the help of such an oracle it would be possible to sample uniformly the optimal assignments of  $F$ , by computing these marginals, setting one unassigned variable according to its marginal, and then proceed in the same way with the simplified formula. A slightly less ambitious, yet still unrealistic, task is to find one optimal configuration (not necessarily uniformly distributed) of  $F$ ; this can be performed if the oracle is able to reveal, for each formula he is questioned about, which of the unassigned variables take the same value in all optimal assignments, and what is this value. Then it is enough to avoid setting incorrectly such a constrained variable to obtain at the end an optimal assignment.

Of course such procedures are not meant as practical algorithms; instead of these fictitious oracles one has to resort to simplified evidences gathered from the current formula to guide the choice of the variable to assign. In Sec. V A we consider algorithms exploiting basic information on the number of occurrences of each variable, and their behavior in the satisfiable regime of random SAT formulas. They are turned into complete algorithms by allowing for backtracking the heuristic choices, as explained in V B. Finally in Sec. V C we shall use more refined message-passing sub-procedures to provide the information used in the assignment steps.



### A. Heuristic search: the success-to-failure transition

The first algorithm we consider was introduced and analyzed by Franco and his collaborators [72, 73].

1. If a formula contains a *unit clause* i.e. a clause with a single variable, this clause is satisfied through an appropriate assignment of its unique variable (propagation); If the formula contains no *unit-clause* a variable and its truth value are chosen according to some heuristic rule (free choice). Note that the unit clause propagation corresponds to the obvious answer an oracle would provide on such a formula.
2. Then the clauses in which the assigned variable appears are simplified: satisfied clauses are removed, the other ones are reduced.
3. Resume from step 1.

The procedure will end if one of two conditions is verified:

1. The formula is completely empty (all clauses have been removed), and a solution has been found (SUCCESS).
2. A contradiction is generated from the presence of two opposite unit clauses. The algorithm halts. We do not know if a solution exists and has not been found or if there is no solution (FAILURE).

The simplest example of heuristic is called Unit Clause (UC) and consists in choosing a variable uniformly at random among those that are not yet set, and assigning it to TRUE or FALSE uniformly at random. More sophisticated heuristics can take into account the number of occurrences of each variable and of its negation, the length of the clauses in which each variable appears, or they can set more than one variable at a time. For example, in the Generalized Unit Clause (GUC), the variable is always chosen among those appearing in the shortest clauses.

Numerical experiments and theory show that the results of this procedure applied to random  $k$ -SAT formulas with ratios  $\alpha$  and size  $N$  can be classified in two regimes:

- At low ratio  $\alpha < \alpha_H$  the search procedure finds a solution with positive probability (over the formulas and the random choices of the algorithm) when  $N \rightarrow \infty$ .
- At high ratio  $\alpha > \alpha_H$  the probability of finding a solution vanishes when  $N \rightarrow \infty$ . Notice that  $\alpha_H < \alpha_s$ : solutions do exist in the range  $[\alpha_H, \alpha_s]$  but are not found by this heuristic.

The above algorithm *modifies* the formula as it proceeds; during the execution of the algorithm the current formula will contain clauses of length 2 and 3 (we specialize here to  $k = 3$ -SAT for the sake of simplicity but higher values of  $k$  can be considered). The sub-formulas generated by the search procedure maintain their statistical uniformity (conditioned on the number of clauses of length 2 and 3). Franco and collaborators used this fact to write down differential equations for the evolution of the densities of 2- and 3-clauses as a function of the fraction  $t$  of eliminated variables. We do not reproduce those equations here, see [74] for a pedagogical review. Based on this analysis Frieze and Suen [75] were able to calculate, in the limit of infinite size, the probability of successful search. The outcome for the UC heuristic is

$$\mathcal{P}_{\text{success}}^{(\text{UC})}(\alpha) = \exp \left\{ -\frac{1}{4\sqrt{8/3\alpha - 1}} \arctan \left[ \frac{1}{\sqrt{8/3\alpha - 1}} \right] - \frac{3}{16}\alpha \right\} \quad (30)$$

when  $\alpha < \frac{8}{3}$ , and  $\mathcal{P} = 0$  for larger ratios. The probability  $\mathcal{P}_{\text{success}}$  is, as expected, a decreasing function of  $\alpha$ ; it vanishes in  $\alpha_H = \frac{8}{3}$ . A similar calculation shows that  $\alpha_H \simeq 3.003$  for the GUC heuristic [75].

Franco et al's analysis can be recast in the following terms. Under the operation of the algorithm the original 3-SAT formula is turned into a mixed  $2 + p$ -SAT formula where  $p$  denotes the fraction of the clauses with 3 variables: there are  $N\alpha \cdot (1 - p)$  2-clauses and  $N\alpha p$  3-clauses. As we mentioned earlier the simplicity of the heuristics maintains a statistical uniformity over the formulas with a given value of  $\alpha$  and  $p$ . This constatation motivated the study of the random  $2 + p$ -SAT ensemble by statistical mechanics methods [20, 56], some of the results being later confirmed by the rigorous analysis of [76]. At the heuristic level one expects the existence of a  $p$  dependent satisfiability threshold  $\alpha_s(p)$ , interpolating between the 2-SAT known threshold,  $\alpha_s(p = 0) = 1$ , and the conjectured 3-SAT case,  $\alpha_s(p = 1) \approx 4.267$ . The upperbound  $\alpha_s(p) \leq 1/(1 - p)$  is easily obtained: for the mixed formula to be satisfiable, necessarily the sub-formula obtained by retaining only the clauses of length 2 must be satisfiable as well. In fact this bound is tight for all values of  $p \in [0, 2/5]$ . During the execution of the algorithm the ratio  $\alpha$  and the fraction  $p$  are 'dynamical' parameters, changing with the fraction  $t = T/N$  of variables assigned by the algorithm. They define the coordinates

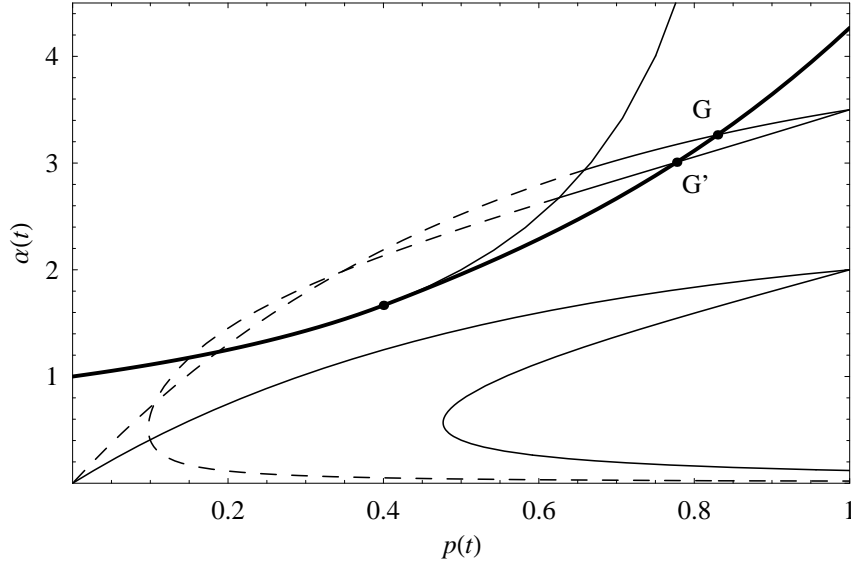


FIG. 7: Trajectories generated by heuristic search acting on 3-SAT for  $\alpha = 2$  and  $\alpha = 3.5$ . For all heuristics, the starting point is on the  $p = 1$  axis, with the initial value of  $\alpha$  as ordinate. The curves that end at the origin correspond to UC, those ending on the  $p = 1$  axis correspond to GUC. The thick line represents the satisfiability threshold: the part on the left of the critical point  $(2/5, 5/3)$  is exact and coincides with the contradiction line, where contradictions are generated with high probability, of equation  $\alpha = 1/(1 - p)$ , and which is plotted for larger values of  $p$  as well; the part on the right of the critical point is only a sketch. When the trajectories hit the satisfiability threshold, at points G for UC and G' for GUC, they enter a region in which massive backtracking takes place, and the trajectory represents the evolution *prior* to backtracking. The dashed part of the curves is “unphysical”, i.e. the trajectories stop when the contradiction curve is reached.

of the representative point of the instance at ‘time’  $t$  in the  $(p, \alpha)$  plane of Figure 7. The motion of the representative point defines the search trajectory of the algorithm. Trajectories start from the point of coordinates  $p(0) = 1, \alpha(0) = \alpha$  and end up on the  $\alpha = 0$  axis when a solution is found. The probability of success is positive as long as the 2-SAT subformula is satisfiable, that is, as long as  $\alpha \cdot (1 - p) < 1$ . In other words success is possible provided the trajectory does not cross the contradiction line  $\alpha = 1/(1 - p)$  (Figure 7). The largest initial ratio  $\alpha$  such that no crossing occurs defines  $\alpha_H$ . Notice that the search trajectory is a stochastic object. However Franco has shown that the deviations from its average locus in the plane vanish in the  $N \rightarrow \infty$  limit (concentration phenomenon). Large deviations from the typical behavior can be calculated e.g. to estimate the probability of success above  $\alpha_H$  [77].

The precise form of  $\mathcal{P}_{\text{success}}$  and the value  $\alpha_H$  of the ratio where it vanishes are specific to the heuristic considered (UC in (30)). However the behavior of the probability close to  $\alpha_H$  is largely independent of the heuristic (provided it preserves the uniformity of the subformulas generated):

$$\ln \mathcal{P}_{\text{success}}(\alpha = \alpha_H(1 - \lambda)) \sim -\lambda^{-1/2}. \quad (31)$$

This universality can loosely be interpreted by observing that for  $\alpha$  close to  $\alpha_H$  the trajectory will pass very close to the contradiction curve  $\alpha \cdot (1 - p) = 1$ , which characterizes the locus of the points where the probability that a variable is assigned by the heuristics  $H$  vanishes (and all the variables are assigned by Unit Propagation). The value of  $\alpha_H$  depend on the “shape” of the trajectory far from this curve, and will therefore depend on the heuristics, but the probability of success (i.e. of avoiding the contradiction curve) for values of  $\alpha$  close to  $\alpha_H$  will only depend on the local behavior of the trajectory close to the contradiction curve, a region where most variables are assigned through Unit Propagation and not sensitive to the heuristics.

The finite-size corrections to equation (30) are also universal (i.e. independent on the heuristics):

$$\ln \mathcal{P}_{\text{success}}(\alpha = \alpha_H(1 - \lambda), N) \sim -N^{1/6} \mathcal{F}(\lambda N^{1/3}), \quad (32)$$

where  $\mathcal{F}$  is a universal scaling function which can be exactly expressed in terms of the Airy function [78]. This result indicates that right at  $\alpha_H$  the probability of success decreases as a stretched exponential  $\sim \exp(-cst N^{1/6})$ .

The exponent  $\frac{1}{3}$  suggests that the critical scaling of  $\mathcal{P}$  is related to random graphs. After  $T = tN$  steps of the procedure, the sub-formula will consists of  $C_3, C_2$  and  $C_1$  clauses of length 3, 2 and 1 respectively (notice that these are *extensive*, i.e.  $O(N)$  quantities). We can represent the clauses of length 1 and 2 (which are the relevant ones to

understand the generation of contradictions) as an oriented graph  $\mathcal{G}$  in the following way. We will have a vertex for each literal, and represent 1-clauses by “marking” the literal appearing in each; a 2-clause will be represented by two directed edges, corresponding to the two implications equivalent to the clause (for example,  $x_1 \vee \bar{x}_2$  is represented by the directed edges  $\bar{x}_1 \rightarrow \bar{x}_2$  and  $x_2 \rightarrow x_1$ ). The average out-degree of the vertices in the graph is  $\gamma = C_2/(N - T) = \alpha(t)(1 - p(t))$ .

What is the effect of the algorithm on  $\mathcal{G}$ ? The algorithm will proceed in “rounds”: a variable is set by the heuristics, and a series of Unit Propagations are performed until no more unit clauses are left, at which point a new round starts. Notice that during a round, extensive quantities as  $C_1, C_2, C_3$  are likely to vary by bounded amounts and  $\gamma$  to vary by  $O(\frac{1}{N})$  (this is the very reason that guarantees that these quantities are concentrated around their mean). At each step of Unit Propagation, a marked literal (say  $x$ ) is assigned and removed from  $\mathcal{G}$ , together with all the edges connected to it, and the “descendants” of  $x$  (i.e. the literals at the end of outgoing edges) are marked. Also  $\bar{x}$  is removed together with its edges, but its descendants are not marked. Therefore, the marked vertices “diffuse” in a connected component of  $\mathcal{G}$  following directed edges. Moreover, at each step new edges corresponding to clauses of length 3 that get simplified into clauses of length 2 are added to the graph.

When  $\gamma > 1$ ,  $\mathcal{G}$  undergoes a directed percolation transition, and a giant component of size  $O(N)$  appears, in which it is possible to go from any vertex to any other vertex by following a directed path. When this happens, there is a finite probability that two opposite literals  $x$  and  $\bar{x}$  can be reached from some other literal  $y$  following a directed path. If  $\bar{y}$  is selected by Unit Propagation, at some time both  $x$  and  $\bar{x}$  will be marked, and this corresponds to a contradiction. This simple argument explains more than just the condition  $\gamma = \alpha \cdot (1 - p) = 1$  for the failure of the heuristic search. It can also be used to explain the the exponent  $\frac{1}{6}$  in the scaling (32), see [78, 79] for more details.

## B. Backtrack-based search: the Davis-Putnam-Loveland-Logeman procedure

The heuristic search procedure of the previous Section can be easily turned into a complete procedure for finding solutions or proving that formulas are not satisfiable. When a contradiction is found the algorithm now backtracks to the last assigned variable (by the heuristic; unit clause propagations are merely consequences of previous assignments), invert it, and the search resumes. If another contradiction is found the algorithm backtracks to the last-but-one assigned variable and so on. The algorithm stops either if a solution is found or all possible backtracks have been unsuccessful and a proof of unsatisfiability is obtained. This algorithm was proposed by Davis, Putnam, Loveland and Logemann and is referred to as DPLL in the following.

The history of the search process can be represented by a search tree, where the nodes represent the variables assigned, and the descending edges their values (Figure 8). The leaves of the tree correspond to solutions (S), or to contradictions (C). The analysis of the  $\alpha < \alpha_H$  regime in the previous Section leads us to the conclusion that search trees look like Figure 8A at small ratios<sup>4</sup>.

For ratios  $\alpha > \alpha_H$  DPLL is very likely to find a contradiction. Backtracking enters into play, and is responsible for the drastic slowing down of the algorithm. The success-to-failure transition takes place in the non-backtracking algorithm into a polynomial-to-exponential transition in DPLL. The question is to compute the growth exponent of the average tree size,  $T \sim e^{N\tau(\alpha)}$ , as a function of the ratio  $\alpha$ .

### 1. Exponential regime: Unsatisfiable formulas

Consider first the case of unsatisfiable formulas ( $\alpha > \alpha_s$ ) where all leaves carry contradictions after DPLL halts (Figure 8B). DPLL builds the tree in a sequential manner, adding nodes and edges one after the other, and completing branches through backtracking steps. We can think of the same search tree built in a parallel way [80]. At time (depth  $T$ ) our tree is composed of  $L(T) \leq 2^T$  branches, each carrying a partial assignment over  $T$  variables. Step  $T$  consists in assigning one more variable to each branch, according to DPLL rules, that is, through unit-propagation or the heuristic rule. In the latter case we will speak of a splitting event, as two branches will emerge from this node, corresponding to the two possible values of the variable assigned. The possible consequences of this assignment are the emergence of a contradiction (which put an end to the branch), or the simplification of the attached formulas (the branch keeps growing).

The number of branches  $L(T)$  is a stochastic variable. Its average value can be calculated as follows [81]. Let us define the average number  $L(\bar{C}; T)$  of branches of depth  $T$  which bear a formula containing  $C_3$  (resp.  $C_2, C_1$ )

---

<sup>4</sup> A small amount of backtracking may be necessary to find the solution since  $\mathcal{P}_{\text{success}} < 1$  [75], but the overall picture of a single branch is not qualitatively affected.

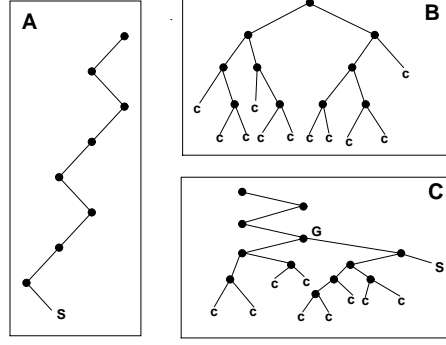


FIG. 8: Search trees generated by DPLL: **A.** linear, satisfiable ( $\alpha < \alpha_H$ ); **B.** exponential, unsatisfiable ( $\alpha > \alpha_c$ ). **C.** exponential, satisfiable ( $\alpha_H < \alpha < \alpha_c$ ); Leaves are marked with S (solutions) or C (contradictions). G is the highest node to which DPLL backtracks, see Figure 7.

equations of length 3 (resp. 2,1), with  $\vec{C} = (C_1, C_2, C_3)$  Initially  $L(\vec{C}; 0) = 1$  for  $\vec{C} = (0, 0, \alpha N)$ , 0 otherwise. We shall call  $M(\vec{C}', \vec{C}; T)$  the average number of branches described by  $\vec{C}'$  generated from a  $\vec{C}$  branch once the  $T^{th}$  variable is assigned [79, 80]. We have  $0 \leq M \leq 2$ , the extreme values corresponding to a contradiction and to a split respectively. We claim that

$$L(\vec{C}'; T+1) = \sum_{\vec{C}} M(\vec{C}', \vec{C}; T) L(\vec{C}; T). \quad (33)$$

Evolution equation (33) could look like somewhat suspicious at first sight due to its similarity with the approximation we have sketched in Sec. IV B for the analysis of PRWSAT. Yet, thanks to the linearity of expectation, the correlations between the branches (or better, the instances carried by the branches) do not matter as far as the average number of branches is concerned.

For large  $N$  we expect that the number of alive (not hit by a contradiction) branches grows exponentially with the depth, or, equivalently,

$$\sum_{C_1, C_2, C_3} L(C_1, C_2, C_3; T) \sim e^{N \lambda(t) + o(N)} \quad (34)$$

The argument of the exponential,  $\lambda(t)$ , can be found using partial differential equation techniques generalizing the ordinary differential equation techniques of a single branch in the absence of backtracking (Section V A). Details can be found in [81]. The outcome is that  $\lambda(t)$  is a function growing from  $\lambda = 0$  at  $t = 0$ , reaching a maximum value  $\lambda_M$  for some depth  $t_M$ , and decreasing at larger depths.  $t_M$  is the depth in the tree of Figure 8B where most contradictions are found; the number of contradiction leaves is, to exponential order,  $e^{N \lambda_M}$ . We conclude that the logarithm of the average size of the tree we were looking for is

$$\tau = \lambda_M. \quad (35)$$

For large  $\alpha \gg \alpha_s$  one finds  $\tau = O(1/\alpha)$ , in agreement with the asymptotic scaling of [82]. The calculation can be extended to higher values of  $k$ .

## 2. Exponential regime: Satisfiable formulas

The above calculation holds for the unsatisfiable, exponential phase. How can we understand the satisfiable but exponential regime  $\alpha_H < \alpha < \alpha_s$ ? The resolution trajectory crosses the SAT/UNSAT critical line  $\alpha_s(p)$  at some point G shown in Figure 7. Immediately after G the instance left by DPLL is unsatisfiable. A subtree with all its leaves carrying contradictions will develop below G (Figure 8C). The size  $\tau^G$  of this subtree can be easily calculated from the above theory from the knowledge of the coordinates  $(p_G, \alpha_G)$  of G. Once this subtree has been built DPLL backtracks to G, flips the attached variable and will finally end up with a solution. Hence the (log of the) number of splits necessary will be equal to  $\tau = (1 - t_G) \tau_{\text{split}}^G$  [80]. Remark that our calculation gives the logarithm of the average subtree size starting from the typical value of G. Numerical experiments show that the resulting value for  $\tau$  coincides very accurately with the most likely tree size for finding a solution. The reason is that fluctuations in the sizes are mostly due to fluctuations of the highest backtracking point G, that is, of the first part of the search trajectory [77].

### C. Message passing algorithms

According to the thought experiment proposed at the beginning of this Section valuable information could be obtained from the knowledge of the marginal probabilities of variables in the uniform measure over optimal configurations. This is an inference problem in the graphical model associated to the formula. In this field message passing techniques (for instance Belief Propagation, or the min-sum algorithm) are widely used to compute approximately such marginals [46, 48]. These numerical procedures introduce messages on the directed edges of the factor graph representation of the problem (recall the definitions given in Sec. III D), which are iteratively updated, the new value of a message being computed from the old values of the incoming messages (see Fig. 4). When the underlying graph is a tree, the message updates are guaranteed to converge in a finite number of steps, and provide exact results. In the presence of cycles the convergence of these recurrence equations is not guaranteed; they can however be used heuristically, the iterations being repeated until a fixed point has been reached (within a tolerance threshold). Though very few general results on the convergence in presence of loops are known [83] (see also [84] for low  $\alpha$  random SAT formulas) these heuristic procedures are often found to yield good approximation of the marginals on generic factor graph problems.

The interest in this approach for solving random SAT instances was triggered in the statistical mechanics community by the introduction of the Survey Propagation algorithm [21]. Since then several generalizations and reinterpretations of SP have been put forward, see for instance [85–90]. In the following paragraph we present three different message passing procedures, which differ in the nature of the messages passed between nodes, following rather closely the presentation of [47] to which we refer the reader for further details. We then discuss how these procedures have to be interleaved with assignment (decimation) steps in order to constitute a solver algorithm. Finally we shall review results obtained in a particular limit case (large  $\alpha$  satisfiable formulas).

#### 1. Definition of the message-passing algorithms

- Belief Propagation (BP)

For the sake of readability we recall here the recursive equations (26) stated in Sec. III D for the uniform probability measure over the solutions of a tree formula,

$$\begin{aligned} h_{i \rightarrow a} &= \sum_{b \in \partial_+ i(a)} u_{b \rightarrow i} - \sum_{b \in \partial_- i(a)} u_{b \rightarrow i} , \\ u_{a \rightarrow i} &= -\frac{1}{2} \ln \left( 1 - \prod_{j \in \partial a \setminus i} \frac{1 - \tanh h_{j \rightarrow a}}{2} \right) . \end{aligned} \quad (36)$$

where the  $h$  and  $u$ 's messages are reals (positive for  $u$ ), parametrizing the marginal probabilities (beliefs) for the value of a variable in absence of some constraint nodes around it (cf. Eq. (25)). These equations can be used in the heuristic way explained above for any formula, and constitute the BP message-passing equations. Note that in the course of the simplification process the degree of the clauses change, we thus adopt here and in the following the natural convention that sums (resp. products) over empty sets of indices are equal to 0 (resp. 1).

- Warning Propagation (WP)

The above-stated version of the BP equations become ill-defined for an unsatisfiable formula, whether this was the case of the original formula or because of some wrong assignment steps; in particular the normalization constant of Eq. (24) vanishes. A way to cure this problem consists in introducing a fictitious inverse temperature  $\beta$  and deriving the BP equations corresponding to the regularized Gibbs-Boltzmann probability law (20), taking as the energy function the number of unsatisfied constraints. In the limit  $\beta \rightarrow \infty$ , in which the Gibbs-Boltzmann measure concentrates on the optimal assignments, one can single out a part of the information conveyed by the BP equations to obtain the simpler Warning Propagation rules. Indeed the messages  $h, u$  are at leading order proportional to  $\beta$ , with proportionality coefficients we shall denote  $\hat{h}$  and  $\hat{u}$ . These messages are less informative than the ones of BP, yet simpler to handle. One finds indeed that instead of reals the WP messages are integers, more precisely  $\hat{h} \in \mathbb{Z}$  and  $\hat{u} \in \{0, 1\}$ . They obey the following recursive equations (with a structure similar to

the ones of BP),

$$\begin{aligned}\hat{h}_{i \rightarrow a} &= \sum_{b \in \partial_+ i(a)} \hat{u}_{b \rightarrow i} - \sum_{b \in \partial_- i(a)} \hat{u}_{b \rightarrow i} , \\ \hat{u}_{a \rightarrow i} &= \prod_{j \in \partial a \setminus i} \mathbb{I}(\hat{h}_{j \rightarrow a} < 0) ,\end{aligned}\tag{37}$$

where  $\mathbb{I}(E)$  is the indicator function of the event  $E$ . The interpretation of these equations goes as follows.  $\hat{u}_{a \rightarrow i}$  is equal to 1 if in all optimal assignments of the amputated formula in which  $i$  is only constrained by  $a$ ,  $i$  takes the value satisfying  $a$ . This happens if all other variables of clause  $a$  (i.e.  $\partial a \setminus i$ ) are required to take their values unsatisfying  $a$ , hence the form of the right part of (37). In such a case we say that  $a$  sends a warning to variable  $i$ . In the first part of (37), the message  $\hat{h}_{i \rightarrow a}$  sent by a variable to a clause is computed by pondering the number of warnings sent by all other clauses; it will in particular be negative if a majority of clauses requires  $i$  to take the value unsatisfying  $a$ .

- Survey Propagation (SP)

The convergence of BP and WP iterations is not ensured on loopy graphs. In particular the clustering phenomenon described in Sec. III A is likely to spoil the efficiency of these procedures. The Survey Propagation (SP) algorithm introduced in [21] has been designed to deal with these clustered space of configurations. The underlying idea is that the simple iterations (of BP or WP type) remain valid inside each cluster of optimal assignments; for each of these clusters  $\gamma$  and each directed edge of the factor graph one has a message  $h_{i \rightarrow a}^\gamma$  (and  $u_{a \rightarrow i}^\gamma$ ). One introduces on each edge a survey of these messages, defined as their probability distribution with respect to the choice of the clusters. Then some hypotheses are made on the structure of the cluster decomposition in order to write closed equations on the survey. We explicit now this approach in a version adapted to satisfiable instances [47], taking as the basic building block the WP equations. This leads to a rather simple form of the survey. Indeed  $\hat{u}_{a \rightarrow i}$  can only take two values, its probability distribution can thus be parametrized by a single real  $\delta_{a \rightarrow i} \in [0, 1]$ , the probability that  $\hat{u}_{a \rightarrow i} = 1$ . Similarly the survey  $\gamma_{i \rightarrow a}$  is the probability that  $\hat{h}_{i \rightarrow a} < 0$ . The second part of (37) is readily translated in probabilistic terms,

$$\delta_{a \rightarrow i} = \prod_{j \in \partial a \setminus i} \gamma_{j \rightarrow a} .\tag{38}$$

The other part of the recursion takes a slightly more complicated form,

$$\begin{aligned}\gamma_{i \rightarrow a} &= \frac{(1 - \pi_{i \rightarrow a}^-) \pi_{i \rightarrow a}^+}{\pi_{i \rightarrow a}^+ + \pi_{i \rightarrow a}^- - \pi_{i \rightarrow a}^+ \pi_{i \rightarrow a}^-} , \\ \text{with } \begin{cases} \pi_{i \rightarrow a}^+ &= \prod_{b \in \partial_+ i(a)} (1 - \delta_{b \rightarrow i}) \\ \pi_{i \rightarrow a}^- &= \prod_{b \in \partial_- i(a)} (1 - \delta_{b \rightarrow i}) \end{cases} .\end{aligned}\tag{39}$$

In this equation  $\pi_{i \rightarrow a}^+$  (resp.  $\pi_{i \rightarrow a}^-$ ) corresponds to the probability that none of the clauses agreeing (resp. disagreeing) with  $a$  on the value of the literal of  $i$  sends a warning. For  $i$  to be constrained to the value unsatisfying  $a$ , at least one of the clauses of  $\partial_- i(a)$  should send a warning, and none of  $\partial_+ i(a)$ , which explains the form of the numerator of  $\gamma_{i \rightarrow a}$ . The denominator arises from the exclusion of the event that both clauses in  $\partial_+ i(a)$  and  $\partial_- i(a)$  send messages, a contradictory event in this version of SP which is devised for satisfiable formulas.

From the statistical mechanics point of view the SP equations arise from a 1RSB cavity calculation, as sketched in Sec. III D, in the zero temperature limit ( $\beta \rightarrow \infty$ ) and vanishing Parisi parameter  $m$ , these two limits being either taken simultaneously as in [21, 89] or successively [22]. One can thus compute, from the solution of the recursive equations on a single formula, an estimation of its complexity, i.e. the number of its clusters (irrespectively of their sizes). The message passing procedure can also be adapted, at the price of technical complications, to unsatisfiable clustered formulas [89]. Note also that the above SP equations have been shown to correspond to the BP ones in an extended configuration space where variables can take a “joker” value [85, 86], mimicking the variables which are not frozen to a single value in all the assignments of a given cluster. Heuristic interpolations between the BP and SP equations have been studied in [86, 87].

## 2. Exploiting the information

The information provided by these message passing procedures can be exploited in order to solve satisfiability formulas; in the algorithm sketched at the beginning of Sec. V A the heuristic choice of the assigned variable, and its truth value, can be done according to the results of the message passing on the current formula. If BP were an exact inference algorithm, one could choose any unassigned variable, compute its marginal according to Eq. (27), and draw it according to this probability. Of course BP is only an approximate procedure, hence a practical implementation of this idea should privilege the variables with marginal probabilities closest to a deterministic law (i.e. with the largest  $|h_i|$ ), motivated by the intuition that these are the least subject to the approximation errors of BP. Similarly, if the message passing procedure used at each assignment step is WP, one can fix the variable with the largest  $|\hat{h}_i|$  to the value corresponding to the sign of  $\hat{h}_i$ . In the case of SP, the solution of the message passing equations are used to compute, for each unassigned variable  $i$ , a triplet of numbers  $(\gamma_i^+, \gamma_i^-, \gamma_i^0)$  according to

$$\gamma_i^+ = \frac{(1 - \pi_i^+) \pi_i^-}{\pi_i^+ + \pi_i^- - \pi_i^+ \pi_i^-}, \quad \gamma_i^- = \frac{(1 - \pi_i^-) \pi_i^+}{\pi_i^+ + \pi_i^- - \pi_i^+ \pi_i^-}, \quad \gamma_i^0 = 1 - \gamma_i^+ - \gamma_i^-,$$

$$\text{with } \begin{cases} \pi_i^+ = \prod_{a \in \partial_+ i} (1 - \delta_{a \rightarrow i}) \\ \pi_i^- = \prod_{a \in \partial_- i} (1 - \delta_{a \rightarrow i}) \end{cases} \quad (40)$$

$\gamma_i^+$  (resp.  $\gamma_i^-$ ) is interpreted as the fraction of clusters in which  $\sigma_i = +1$  (resp.  $\sigma_i = -1$ ) in all solutions of the cluster, hence  $\gamma_i^0$  corresponds to the clusters in which  $\sigma_i$  can take both values. In the version of [47], one then choose the variable with the largest  $|\gamma_i^+ - \gamma_i^-|$ , and fix it to  $\sigma_i = +1$  (resp.  $\sigma_i = -1$ ) if  $\gamma_i^+ > \gamma_i^-$  (resp.  $\gamma_i^+ < \gamma_i^-$ ). In this way one tries to select an assignment preserving the maximal number of clusters.

Of course many variants of these heuristic rules can be devised; for instance after each message passing computation one can fix a finite fraction of the variables (instead of a single one), allows for some amount of backtracking [91], or increase a soft bias instead of assigning completely a variable [90]. Moreover the tolerance on the level of convergence of the message passing itself can also be adjusted. All these implementation choices will affect the performances of the solver, in particular the maximal value of  $\alpha$  up to which random SAT instances are solved efficiently, and thus makes difficult a precise statement about the limits of these algorithms. In consequence we shall only report the impressive result of [47], which presents an implementation [92] working for random 3-SAT instances up to  $\alpha = 4.24$  (very close to the conjectured satisfiability threshold  $\alpha_s \approx 4.267$ ) for problem sizes as large as  $N = 10^7$ .

The theoretical understanding of these message passing inspired solvers is still poor compared to the algorithms studied in Sec. V A, which use much simpler heuristics in their assignment steps. One difficulty is the description of the residual formula after an extensive number of variables have been assigned; because of the correlations between successive steps of the algorithm this residual formula is not uniformly distributed conditioned on a few dynamical parameters, as was the case with  $(\alpha(t), p(t))$  for the simpler heuristics of Sec. V A. One version of BP guided decimation could however be studied analytically in [93], by means of an analysis of the thought experiment discussed at the beginning of Sec. V. The study of another simple message passing algorithm is presented in the next paragraph.

## 3. Warning Propagation on dense random formulas

Feige proved in [94] a remarkable connection between the *worst-case* complexity of approximation problems and the structure of *random* 3-SAT at large (but independent of  $N$ ) values of the ratio  $\alpha$ . He introduced the following hardness hypothesis for random 3-SAT formulas:

**Hypothesis 1:** *Even if  $\alpha$  is arbitrarily large (but independent of  $N$ ), there is no polynomial time algorithm that on most 3-SAT formulas outputs UNSAT, and always outputs SAT on a 3-SAT formula that is satisfiable.*

and used it to derive hardness of approximation results for various computational problems. As we have seen these instances are typically unsatisfiable; the problem of interest is thus to recognize efficiently the rare satisfiable instances of the distribution.

A variant of this problem was studied in [95], where WP was proven to be effective in finding solutions of dense planted random formulas (the planted distribution is the uniform distribution conditioned on being satisfied by a given assignment). More precisely, [95] proves that for  $\alpha$  large enough (but independent of  $N$ ), the following holds with probability  $1 - e^{-O(\alpha)}$ :

1. WP converges after at most  $O(\ln N)$  iterations.

2. If a variable  $i$  has  $\hat{h}_i \neq 0$ , then the sign of  $\hat{h}_i$  is equal to the value of  $\sigma_i$  in the planted assignment. The number of such variables is bigger than  $N(1 - e^{-O(\alpha)})$  (i.e. almost all variables can be reconstructed from the values of  $\hat{h}_i$ ).
3. Once these variables are fixed to their correct assignments, the remaining formula can be satisfied in time  $O(N)$  (in fact, it is a tree formula).

On the basis of non-rigorous statistical mechanics methods, these results were argued in [96] to remain true when the planted distribution is replaced by the uniform distribution conditioned on being satisfiable. In other words by iterating WP for a number of iterations bigger than  $O(\ln N)$  one is able to detect the rare satisfiable instances at large  $\alpha$ . The argument is based on the similarity of structure between the two distributions at large  $\alpha$ , namely the existence of a single, small cluster of solutions where almost all variables are frozen to a given value. This correspondence between the two distributions of instances was proven rigorously in [97], where it was also shown that a related polynomial algorithm succeeds with high probability in finding solutions of the satisfiable distribution of large enough density  $\alpha$ .

These results indicate that a stronger form of hypothesis 1, obtained by replacing *always* with *with probability  $p$*  (with respect to the uniform distribution over the formulas and possibly to some randomness built in the algorithm), is wrong for any  $p < 1$ . However, the validity of hypothesis 1 is still unknown for random 3-SAT instances. Nevertheless, this result is interesting because it is one of the rare cases in which the performances of a message-passing algorithm could be analyzed in full detail.

## VI. CONCLUSION

This review was mainly dedicated to the random  $k$ -Satisfiability and  $k$ -Xor-Satisfiability problems; the approach and results we presented however extend to other random decision problems, in particular random graph  $q$ -coloring. This problem consists in deciding whether each vertex of a graph can be assigned one out of  $q$  possible colors, without giving the same color to the two extremities of an edge. When input graphs are randomly drawn from Erdős-Renyi (ER) ensemble  $G(N, p = c/N)$  a phase diagram similar to the one of  $k$ -SAT (Section III) is obtained. There exists a colorable/uncolorable phase transition for some critical average degree  $c_s(q)$ , with for instance  $c_s(3) \simeq 4.69$  [98]. The colorable phase also exhibits the clustering and condensation transitions [99] we explained on the example of the  $k$ -Satisfiability. Actually what seems to matter here is rather the structure of inputs and the symmetry properties of the decision problem rather than its specific details. All the above considered input models share a common, underlying ER random graph structure. From this point of view it would be interesting to ‘escape’ from the ER ensemble and consider more structured graphs e.g. embedded in a low dimensional space.

To what extent the similarity between phase diagrams correspond to similar behaviour in terms of hardness of resolution is an open question. Consider the case of rare satisfiable instances for the random  $k$ -SAT and  $k$ -XORSAT well above their sat/unsat thresholds (Section V). Both problems share very similar statistical features. However, while a simple message-passing algorithm allows one to easily find a (the) solution for the  $k$ -SAT problem this algorithm is inefficient for random  $k$ -XORSAT. Actually the local or decimation-based algorithms of Sections IV and V are efficient to find solution to rare satisfiable instances of random  $k$ -SAT [100], but none of them works for random  $k$ -XORSAT (while the problem is in P!). This example raises the important question of the relationship between the statistical properties of solutions (or quasi-solutions) encoded in the phase diagram and the (average) computational hardness. Very little is known about this crucial point; on intuitive grounds one could expect the clustering phenomenon to prevent an efficient solving of formulas by local search algorithms of the random walk type. This is indeed true for a particular class of stochastic processes [101], those which respect the so-called detailed balance conditions. This connection between clustering and hardness of resolution for local search algorithms is much less obvious when the detailed balance conditions are not respected, which is the case for most of the efficient variants of PRWSAT.

- 
- [1] M. Mézard, G. Parisi, and M. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).
  - [2] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity* (Dover, New York, 1998).
  - [3] Y. Fu and P. W. Anderson, Journal of Physics A: Mathematical and General **19**, 1605 (1986).
  - [4] D. Mitchell, B. Selman, and H. Levesque (1992), no. 459 in Proceedings of the Tenth National Conference on Artificial Intelligence.
  - [5] J. Hertz, A. Krogh, and R. Palmer, *Introduction to the theory of neural computation*, Santa Fe Institute Studies in the Science of Complexity (Addison-Wesley, Redwood city (CA), 1991).



- [6] T. Cover, IEEE Transactions on Electronic Computers **14**, 326 (1965).
- [7] S. Janson, T. Luczak, and A. Rucinski, *Random graphs* (John Wiley and Sons, New York, 2000).
- [8] E. Friedgut, Journal of the American Mathematical Society **12**, 1017 (1999).
- [9] O. Dubois, Theoret. Comput. Sci. **265**, 187 (2001).
- [10] J. Franco, Theoret. Comput. Sci. **265**, 147 (2001).
- [11] D. Achlioptas and Y. Peres, Journal of the American Mathematical Society **17**, 947 (2004).
- [12] *Chapter random sat, this volume.*
- [13] N. Alon and J. Spencer, *The probabilistic method* (John Wiley and sons, New York, 2000).
- [14] A. Dembo and O. Zeitouni, *Large deviations. Theory and applications* (Springer, Berlin, 1998).
- [15] W. Krauth and M. Mezard, J. Physique **50**, 3057 (1989).
- [16] S. K. Ma, *Statistical Mechanics* (World Scientific, Singapore, 1985).
- [17] K. Huang, *Statistical Mechanics* (John Wiley and Sons, New York, 1990).
- [18] A. Broder, A. Frieze, and E. Upfal (1993), no. 322 in Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms.
- [19] R. Monasson and R. Zecchina, Phys. Rev. E **56**, 1357 (1997).
- [20] G. Biroli, R. Monasson, and M. Weigt, Eur. Phys. J. B **14**, 551 (2000).
- [21] M. Mézard and R. Zecchina, Phys. Rev. E **66**, 056126 (2002).
- [22] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborova, Proceedings of the National Academy of Sciences **104**, 10318 (2007), <http://www.pnas.org/cgi/reprint/104/25/10318.pdf>.
- [23] R. Monasson and D. O’Kane, Europhysics Letters **27**, 85 (1994).
- [24] T. R. Kirkpatrick and D. Thirumalai, Phys. Rev. B **36**, 5388 (1987).
- [25] M. Talagrand, *Spin glasses: a challenge for mathematicians* (Springer, Berlin, 2003).
- [26] D. Panchenko and M. Talagrand, Probab. Theory Relat. Fields **130**, 319 (2004).
- [27] S. Franz and M. Leone, J. Stat. Phys. **111**, 535 (2003).
- [28] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, J. Stat. Phys. **111**, 505 (2003).
- [29] S. Cocco, O. Dubois, J. Mandler, and R. Monasson, Phys. Rev. Lett. **90**, 047205 (2003).
- [30] M. Mézard, T. Mora, and R. Zecchina, Physical Review Letters **94**, 197205 (pages 4) (2005).
- [31] H. Daudé, M. Mézard, T. Mora, and R. Zecchina (2005), [arXiv:cond-mat/0506053](https://arxiv.org/abs/cond-mat/0506053).
- [32] D. Achlioptas and F. Ricci-Tersenghi, Proceedings of the thirty-eighth annual ACM symposium on Theory of computing (2006), [arXiv:cs.CC/0611052](https://arxiv.org/abs/cs.CC/0611052).
- [33] F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Phys. Rev. E **63**, 026702 (2001).
- [34] B. Pittel, J. Spencer, and N. Wormald, J. Comb. Theory, Ser. B **67**, 111 (1996).
- [35] T. Kurtz, J. Appl. Probab. **7**, 49 (1970).
- [36] A. Montanari and G. Semerjian, J. Stat. Phys. **124**, 103 (2006).
- [37] T. Mora and M. Mézard, Journal of Statistical Mechanics: Theory and Experiment **2006**, P10007 (2006).
- [38] S. Mertens, M. Mézard, and R. Zecchina, Random Struct. Algorithms **28**, 340 (2006).
- [39] M. Mézard, M. Palassini, and O. Rivoire, Physical Review Letters **95**, 200202 (pages 4) (2005).
- [40] A. Montanari, G. Parisi, and F. Ricci-Tersenghi, Journal of Physics A: Mathematical and General **37**, 2073 (2004).
- [41] T. Mora and L. Zdeborova (2007), [arXiv:0710.3804](https://arxiv.org/abs/0710.3804).
- [42] G. Semerjian, J.Stat.Phys. **130**, 251 (2008).
- [43] R. Monasson, Journal of Physics A: Mathematical and General **31**, 513 (1998).
- [44] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
- [45] M. Mézard and G. Parisi, J. Stat. Phys. **111**, 1 (2003).
- [46] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, IEEE Trans. Inf. Theory **47**, 498 (2001).
- [47] A. Braunstein, M. Mézard, and R. Zecchina, Random Struct. Algorithms **27**, 201 (2005).
- [48] J. S. Yedidia, W. T. Freeman, and Y. Weiss, Advances in Neural Information Processing Systems **13**, 689 (2001).
- [49] J. S. Yedidia, W. T. Freeman, and Y. Weiss, in *Exploring Artificial Intelligence in the New Millennium* (2003), p. 239.
- [50] W. Fernandez de la Vega, Theor. Comput. Sci. **265**, 131 (2001).
- [51] B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson, Random Struct. Algorithms **18**, 201 (2001).
- [52] A. Amraoui, A. Montanari, T. Richardson, and R. Urbanke, [arXiv:cs.IT/0406050](https://arxiv.org/abs/cs.IT/0406050) (2004).
- [53] A. Dembo and A. Montanari, [arXiv:math.PR/0702007](https://arxiv.org/abs/math.PR/0702007) (2007).
- [54] D. B. Wilson, Random Struct. Algorithms **21**, 182 (2002).
- [55] S. Kirkpatrick and B. Selman, Science **264**, 1297 (1994).
- [56] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Random Struct. Algorithms **15**, 414 (1999).
- [57] P. De Gregorio, A. Lawlor, P. Bradley, and K. Dawson, PNAS **102**, 5669 (2005).
- [58] L. Cugliandolo, in *Slow relaxations and nonequilibrium dynamics in condensed matter*, edited by J. L. Barrat, M. Feigelman, J. Kurchan, and J. Dalibard (Springer-Verlag, Les Houches, France, 2003).
- [59] C. Papadimitriou, in *Proceedings of the 32th Annual Symposium on Foundations of Computer Science* (1991), pp. 163–169.
- [60] R. Motwani and P. Ravaghan, *Randomized algorithms* (Cambridge University Press, Cambridge, 1995).
- [61] U. Schöning, Algorithmica **32**, 615 (2002), ISSN 0178-4617 (print), 1432-0541 (electronic).
- [62] S. Baumer and R. Schuler, Lecture Notes in Computer Science **2919**, 150 (2004).
- [63] G. Semerjian and R. Monasson, Phys. Rev. E **67**, 066103 (2003).
- [64] W. Barthel, A. K. Hartmann, and M. Weigt, Phys. Rev. E **67**, 066104 (2003).
- [65] T. M. Liggett, *Interacting particle systems* (Springer, Berlin, 1985).

- [66] M. Alekhnovich and E. Ben-Sasson, SIAM Journal on Computing **36**, 1248 (2006).
- [67] B. Selman, H. A. Kautz, and B. Cohen, in *Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI'94)* (Seattle, 1994), pp. 337–343.
- [68] D. McAllester, B. Selman, and H. Kautz, in *Proceedings of the Fourteenth National Conference on Artificial Intelligence (AAAI'97)* (Providence, Rhode Island, 1997), pp. 321–326.
- [69] S. Seitz, M. Alava, and P. Orponen, Journal of Statistical Mechanics: Theory and Experiment **2005**, P06006 (2005).
- [70] J. Ardelius and E. Aurell, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **74**, 037702 (pages 4) (2006).
- [71] M. Alava, J. Ardelius, E. Aurell, P. Kaski, S. Krishnamurthy, P. Orponen, and S. Seitz (2007), [arXiv:0711.4902](#).
- [72] M.-T. Chao and J. Franco, SIAM J. Comput. **15**, 1106 (1986).
- [73] M.-T. Chao and J. Franco, Inf. Sci. **51**, 289 (1990).
- [74] D. Achlioptas, Theor. Comput. Sci. **265**, 159 (2001).
- [75] A. Frieze and S. Suen, J. Algorithms **20**, 312 (1996).
- [76] D. Achlioptas, L. Kirousis, E. Kranakis, and D. Krizanc, Theor. Comput. Sci. **265**, 109 (2001).
- [77] S. Cocco and R. Monasson, Ann. Math. Artif. Intell. **43**, 153 (2005).
- [78] C. Deroulers and R. Monasson, Europhysics Letters **68**, 153 (2004).
- [79] R. Monasson, in *Complex Systems*, edited by J. P. Bouchaud, M. Mézard, and J. Dalibard (Elsevier, Les Houches, France, 2007).
- [80] S. Cocco and R. Monasson, Phys. Rev. Lett. **86**, 1654 (2001).
- [81] R. Monasson, *A generating function method for the average-case analysis of DPLL.*, Lecture Notes in Computer Science 3624, 402–413 (2005). (2005).
- [82] P. Beame, R. Karp, T. Pitassi, and M. Saks, SIAM Journal of Computing **31**, 1048 (2002).
- [83] S. Tatikonda and M. Jordan, in *Proc. Uncertainty in Artificial Intell.* (2002), vol. 18, pp. 493–500.
- [84] A. Montanari and D. Shah, in *SODA* (2007), pp. 1255–1264.
- [85] A. Braunstein and R. Zecchina, Journal of Statistical Mechanics: Theory and Experiment **2004**, P06007 (2004).
- [86] E. Maneva, E. Mossel, and M. J. Wainwright, in *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms* (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005), pp. 1089–1098, ISBN 0-89871-585-7.
- [87] E. Aurell, U. Gordon, and S. Kirkpatrick, in *NIPS* (2004).
- [88] G. Parisi (2003), [arXiv:cs.CC/0301015](#).
- [89] D. Battaglia, M. Kolář, and R. Zecchina, Phys. Rev. E **70**, 036107 (2004).
- [90] J. Chavas, C. Furtlehner, M. Mézard, and R. Zecchina, Journal of Statistical Mechanics: Theory and Experiment **2005**, P11016 (2005).
- [91] G. Parisi (2003), [arXiv:cond-mat/0308510](#).
- [92] URL <http://www.ictp.trieste.it/~zecchina/SP>.
- [93] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian (2007), [arXiv:0709.1667](#), to be published in the Proceedings of the 45th Allerton Conference (2007).
- [94] U. Feige, in *STOC* (2002), pp. 534–543.
- [95] U. Feige, E. Mossel, and D. Vilenchik, *Complete convergence of message passing algorithms for some satisfiability problems.*, Lecture Notes in Computer Science 4110, 339–350 (2006). (2006).
- [96] F. Altarelli, R. Monasson, and F. Zamponi, Journal of Physics A: Mathematical and Theoretical **40**, 867 (2007).
- [97] A. Coja-Oghlan, M. Krivelevich, and D. Vilenchik, *Why almost all k-cnf formulas are easy*, to appear (2007).
- [98] F. Krzakala, A. Pagnani, and M. Weigt, Phys. Rev. E **70**, 046705 (2004).
- [99] L. Zdeborová and F. Krzakala, Physical Review E (Statistical, Nonlinear, and Soft Matter Physics) **76**, 031131 (pages 29) (2007).
- [100] W. Barthel, A. K. Hartmann, M. Leone, F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, Phys. Rev. Lett. **88**, 188701 (2002).
- [101] A. Montanari and G. Semerjian, J. Stat. Phys. **125**, 23 (2006).